

46  
10/6/19-685 Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 18:54:54 ; Search time 7554 Seconds  
(without alignments)  
11905.382 Million cell updates/sec

Title: US-10-619-685-1  
Perfect score: 2468  
Sequence: 1 aaaaaattttcaatcac.....aaaaaaaaaaaaaaaaaaaaa 2468

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues  
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.4	22.5	565	CF323765	CF323765 HDN--04-L
2	551	22.3	753	AG021802	AG021802 Oryza sat
3	542.4	22.0	564	CF323815	CF323815 HDN--04-N
4	494.8	20.0	706	CL833850	CL833850 OR Cba005
5	463	18.8	646	CA498690	CA498690 WHE3246 E
6	446.2	18.1	604	CA031067	CA031067 HX08017i
7	446	18.1	627	BU037770	BU037770 946142A06
8	429.2	17.4	1076	CA254411	CA254411 SCBFL411
9	411.8	16.7	667	CA298359	CA298359 SCRFPL803
10	410.4	16.6	617	BG487823	BG487823 FM1_73 H0
11	404.2	16.4	807	CK193855	CK193855 FGAS00227
12	402.4	16.3	548	BG604893	BG604893 WHE0944 F
13	398.4	16.1	1986	CNS02AG0	EX827777 Arabidops
14	395.6	16.0	2195	CNS09YQA	EX831110 Arabidops
15	394	16.0	699	AL506716	AL506716 AL506716
16	391.8	15.9	675	CA235572	CA235572 SCACFL501
17	391.2	15.9	450	CF330979	CF330979 NACL--06-
18	385	15.6	526	CA251300	CA251300 SCOSFL112
19	376	15.2	795	CD122996	CD122996 BES182410
20	375.8	15.2	594	CD979653	CD979653 QAG5D03.X
21	375	15.2	545	CA232369	CA232369 SCRFPL306
22	365.4	14.8	641	AV833315	AV833315 AV833315
23	359.8	14.5	494	BJ224222	BJ224222 BJ224222
24	356.8	14.5	719	AY108867	AY108867 Zea mays

25	350.2	14.2	592	2	BF484964	BF484964 WHE2333.F
26	349.4	14.2	635	6	CF038174	CF038174 QCH17F09.
27	347.4	14.1	611	2	BE511523	BE511523 946061D08
28	341.2	13.8	510	4	BG465723	BG465723 RH122_47
29	340.8	13.8	507	6	CA831022	CA831022 1117014E0
30	322.6	13.1	517	6	CA249488	CA249488 SCRFPL111
31	320.4	13.0	504	6	CD979546	CD979546 QAG4305.X
32	319.4	12.9	744	4	BG646008	BG646008 ESTS07627
33	314.2	12.7	758	5	BQ990215	BQ990215 QCF19M11.
34	314	12.7	569	4	BJ478725	BJ478725 BJ478725
35	301	12.2	1583	3	AY104938	AY104938 Zea mays
36	297.6	12.1	400	2	BF484550	BF484550 WHE2324.H
37	297	12.0	562	4	BG158382	BG158382 FM1_59.F1
38	291.8	11.8	510	4	BM097574	BM097574 EBem04_SQ
39	287.8	11.7	584	1	AJ500791	AJ500791 AJ500791
40	283.2	11.5	412	4	BG241888	BG241888 RH122_51
41	282.2	11.4	802	7	CF922842	CF922842 gmrhrw24
42	282	11.4	707	5	BU004982	BU004982 QGG6M05.Y
43	280.6	11.4	550	5	BQ608748	BQ608748 BRY_4659
44	272.8	11.1	533	5	BQ620280	BQ620280 talr1173G
45	266	10.8	939	8	BZ795282	BZ795282 PUFAN64TD

ALIGNMENTS

RESULT 1  
CF323765  
LOCUS  
DEFINITION  
CF323765 565 bp mRNA linear EST 18-AUG-2003  
HDN--04-L11.g1 OSHDAC1-overexpressing transgenic rice lambda phage  
CDNA library II (HDN) Oryza sativa (japonica cultivar-group) CDNA  
clone HDN--04-L11, mRNA sequence.  
ACCESSION  
CF323765  
VERSION  
CF323765.1 GI:33795792  
KEYWORDS  
EST.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1 (bases 1 to 565)  
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.  
FEATURES  
source  
location/Qualifiers  
1..565  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="HDN--04-L11"  
/tissue\_type="callus"  
/dev\_stage="proliferated callus on 2N6 media for 2 weeks"  
/lab\_host="E.coli SOLR"  
/clone\_lib="OSHDAC1-overexpressing transgenic rice lambda  
phage CDNA library II (HDN)"  
/notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:  
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at  
5' end with EcoRI and 3' end with XhoI site. mRNA was  
derived from rice Histone Deacetylase overexpression  
line."  
ORIGIN  
Query Match 22.5%; Score 555.4; DB 6; Length 565;  
Best Local Similarity 99.8%; Pred. No. 8.9e-132;

Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 755 TTTGAAATTTTGCATTCCTTATGTGAGGTGCTTGGCATAGATTTCAGCAACACAGTATA 814  
Db 9 TTTGAAATTTTGCATTCCTTATGTGAGGTGCTTGGCATAGATTTCAGCAACACAGTATA 68

Qy 815 TGAGGTGCACCTTACCTTGGAGCAAGCTGAAGGATCAAGAAACCTTGAGCAACACTTACAAAT 874  
Db 69 TGAGGTGCACCTTACCTTGGAGCAAGCTGAAGGATCAAGAAACCTTGAGCAACACTTACAAAT 128

Qy 875 TGAGAGGCTTCTGCTGCGCAACAGACTCGGTGAATGAAGATCCAAAGGATGCAAAA 934  
Db 129 TGAGAGGCTTCTGCTGCGCAACAGACTCGGTGAATGAAGATCCAAAGGATGCAAAA 188

Qy 935 GCTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCACTTGA 994  
Db 189 GCTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCACTTGA 248

Qy 995 AGAGCAGCAGATGTTTTCAGATATATCAATGCAATATTCCTGCTTCAGAAACAGAGATAAC 1054  
Db 249 AGAGCAGCAGATGTTTTCAGATATATCAATGCAATATTCCTGCTTCAGAAACAGAGATAAC 308

Qy 1055 TGAACCAACACCTCTCCACAGATTTCTTGAATTCGAAATTCGAAATCTGAGGTGTTAAGGCT 1114  
Db 309 TGAACCAACACCTCTCCACAGATTTCTTGAATTCGAAATTCGAAATCTGAGGTGTTAAGGCT 368

Qy 1115 TGAACCACTGAAGCAAGTAAAGATGCTGTTTAAAGAAAGAGAGACTAGA 1174  
Db 369 TGAACCACTGAAGCAAGTAAAGATGCTGTTTAAAGAAAGAGAGACTAGA 428

Qy 1175 AGAGCATAGAAGCTGCTCATCTTGTGGCGAGGAGGTTATGAGAGGAGTTTAGCAT 1234  
Db 429 AGAGCATAGAAGCTGCTCATCTTGTGGCGAGGAGGTTATGAGAGGAGTTTAGCAT 488

Qy 1235 TGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTGTAACAAATGTAAGCTCA 1294  
Db 489 TGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTGTAACAAATGTAAGCTCA 548

Qy 1295 CATTGCAACAGTGAAG 1311  
Db 549 CATTGCAACAGTGAAG 565

RESULT 2  
AG021802  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) DNA, clone:NC0266.0.102.1A,  
3' flanking sequence of Tos17 insertion in rice strain NC0266,  
genomic survey sequence.

ACCESSION  
AG021802  
VERSION  
AG021802.1  
KEYWORDS  
GSS.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
1  
Miyao.A., Tanaka.K., Murata.K., Sawaki.H., Takeda.S., Abe.K.,  
Shinozuka.Y., Onosato.K. and Hirochika.H.  
Target site specificity of the Tos17 retrotransposon shows a  
preference for insertion within genes and against insertion in  
retrotransposon-rich regions of the genome  
Plant Cell 15 (8), 1771-1780 (2003)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
2  
12897251  
AUTHORS  
Miyao.A., Tanaka.K. and Hirochika.H.  
TITLE  
Direct Submission  
Submitted (25-OCT-1999) Akio Miyao, National Institute of  
Agrobiological Sciences, Molecular Genetics; 2-1-2, Kannondai,  
Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp,  
URL:http://tos.nias.affrc.go.jp/, Tel:81-298-38-7020,

FEATURES  
source  
Location/Qualifiers  
1..753  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/strain="NC0266"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="NC0266.0.102.1A"  
/clone\_lib="PCR product directly amplified from rice  
genomic DNA"  
/note="The 3' end of retrotransposon Tos17 was found  
immediately upstream of this sequence."  
misc\_feature  
1  
/note="This insertion point is base 20049 in AP003452."

ORIGIN  
Query Match 22.3%; Score 551; DB 9; Length 753;  
Best Local Similarity 98.2%; Pred. No. 1.3e-130;  
Matches 554; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1891 GTTTGGACATTTGCCGGAATTCCTATCAAGAAGTTGTTCTTCAATGCCAGTACTCTACGTG 1950  
Db 1 GTTTGGACATTTGCCGGAATTCCTATCAAGAAGTTGTTCTTCAATGCCAGTACTCTACGTG 60

Qy 1951 AGACGGAGACACCTTCGTAAACCTTTTCTCAGATCACACAGGAAACAGTGTCTCGTCCA 2010  
Db 61 AGACGGAGACACCTTCGTAAACCTTTTCTCAGATCACACAGGAAACAGTGTCTCGTCCA 120

Qy 2011 CGCCTGTGCGCCCTTATCACCATAAATGATGAGGATGATGAGAACAGGACTCCGAAGACAT 2070  
Db 121 CGCCTGTGCGCCCTTATCACCATAAATGATGAGGATGATGAGAACAGGACTCCGAAGACAT 180

Qy 2071 TTACAGACACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCAGATGCAATGA 2130  
Db 181 TTACAGACACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCAGATGCAATGA 240

Qy 2131 CTCCCTCTCTGGCCCAACAAGTTTTCAGCAACTCCAGTTTCCCTGTTGTTTACGACAGCCAG 2190  
Db 241 CTCCCTCTCTGGCCCAACAAGTTTTCAGCAACTCCAGTTTCCCTGTTGTTTACGACAGCCAG 300

Qy 2191 AGGTAACATTTGCAGGAGGACATCGACTACTCTTTTGAAGAAAGCGGCTCGCCATCTATC 2250  
Db 301 AGGTAACATTTGCAGGAGGACATCGACTACTCTTTTGAAGAAAGCGGCTCGCCATCTATC 360

Qy 2251 TGGCCAGGCAATGGTTTAACTGTTGATCAATTTATGTACGTAGTTGAAATCTGACTGCA 2310  
Db 361 TGGCCAGGCAATGGTTTAACTGTTGATCAATTTATGTACGTAGTTGAAATCTGACTGCA 420

Qy 2311 TTTTCTTGTGCGTGGCCATTTGCGTATGTTGTCACCAATAGTCGGCTTTCCAGTAGCAC 2370  
Db 421 TTTTCTTGTGCGTGGCCATTTGCGTATGTTGTCACCAATAGTCGGCTTTCCAGTAGCAC 480

Qy 2371 TATTCTGATTTACTGCAATTTGTTTAAATGTTTCTACAAACAGTAAACAGCTCTATACA 2430  
Db 481 TATTCTGATTTACTGCAATTTGTTTAAATGTTTCTACAAACAGTAAACAGCTCTATACA 540

Qy 2431 TTAGCTTGCTCAGCTAAAAA 2454  
Db 541 TTAGCTTGCTCAGCTAACNAGTACA 564

RESULT 3  
CF323815  
LOCUS  
DEFINITION  
HDN--04-N13.g1 OsHDAC1-overexpressing transgenic rice lambda phage  
cDNA library II (HDN) Oryza sativa (japonica cultivar-group) cDNA  
clone HDN--04-N13, mRNA sequence.

ACCESSION  
CF323815  
VERSION  
CF323815.1  
KEYWORDS  
EST.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 564)

Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,

Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm, B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongji University

Yongin, Gyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.myongji.ac.kr.

Location/Qualifiers

1..564

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nackdong"

/db\_xref="taxon:39947"

/clone="HDN-04-N13"

/tissue\_type="callus"

/dev\_stage="proliferated callus on 2N6 media for 2 weeks"

/lab\_host="E.coli SOLR"

/clone\_lib="OshDAC1-overexpressing transgenic rice lambda

phage cDNA library II (HDN)"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:

XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at

5' end with EcoRI and 3' end with XhoI site. mRNA was

derived from rice Histone Deacetylase overexpression

line."

## ORIGIN

Query Match 22.0%; Score 542.4; DB 6; Length 564;  
Best Local Similarity 99.6%; Pred. No. 2e-128;  
Matches 554; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 755 TTGAAATCTTTCGATCTTATGTCAGGTGCTTGGCATAGATTTCACGCAACAGTATA 814

DB 9 TTGAAATCTTTCGATCTTATGTCAGGTGCTTGGCATAGATTTCACGCAACAGTATA 68

QY 815 TGAGGTGCACCTTAGCTTGGACGAAGCTCAAGGATCAAGAGAACCTGAGCAACACTACAAT 874

DB 69 TGAGGTGCACCTTAGCTTGGACGAAGCTCAAGGATCAAGAGAACCTGAGCAACACTACAAT 128

QY 875 TGAGAGGCTTGTCTGTCGCGCAACAGACATGCGTGAATGAAGATCAAGAGGATGCAAAA 934

DB 129 TGAGAGGCTTGTCTGTCGCGCAACAGACATGCGTGAATGAAGATCAAGAGGATGCAAAA 188

QY 935 GCTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCACTTGA 994

DB 189 GCTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCACTTGA 248

QY 995 AGAGCAGCAGATGTTTCAGATATAACATGCAATATTGCTGCTTCAGAACAGAGATAAC 1054

DB 249 AGAGCAGCAGATGTTTCAGATATAACATGCAATATTGCTGCTTCAGAACAGAGATAAC 308

QY 1055 TGAACCAAAACACCTCTCCACAGATTTCTCGAATTTATGCGAATCTGAGGTGTTAAGGCT 1114

DB 309 TGAACCAAAACACCTCTCCACAGATTTCTCGAATTTATGCGAATCTGAGGTGTTAAGGCT 368

QY 1115 TGAACCAACTGAAGCAAGTAAGATGAAGATCTTGTGTTTAAAAAGAAAGCAGAACTAGA 1174

DB 369 TGAACCAACTGAAGCAAGTAAGATGAAGATCTTGTGTTTAAAAAGAAAGCAGAACTAGA 428

QY 1175 AGAGCATA-CAAGCGTCTCATCTTGTTCGCGAGGAAGGTATGCGAGGAGTTTAGCA 1233

DB 429 AGAGCATA-CAAGCGTCTCATCTTGTTCGCGAGGAAGGTATGCGAGGAGTTTAGCA 488

QY 1234 TTGAAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTC 1293

DB 489 TTGAAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTC 548

QY 1294 ACATTGCAACAGTGAA 1309  
DB 549 ACATTGCAACAGTGAA 564

## RESULT 4

CL833850

LOCUS

DEFINITION

OR\_CBA0057B05.f OR\_CBA Oryza rufipogon genomic clone OR\_CBA0057B05

5'- genomic survey sequence.

ACCESSION

CL833850

VERSION

CL833850.1 GI:51079460

KEYWORDS

GSS.

SOURCE

Oryza rufipogon

ORGANISM

Oryza rufipogon

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 706)

Kim, H., Yu, Y., Wissotski, M., Yost, D., Stum, D., Rao, K., Luo, M.,

Jetty, R., Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and

Wing, R.

OMAP Project

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GSC ACC CCA

Plate: 0057 row: B column: 05

Seq primer: TAA TAC GAC TCA TAG GG

Class: BAC ends.

Location/Qualifiers

1..706

/organism="Oryza rufipogon"

/mol\_type="genomic DNA"

/db\_xref="taxon:4529"

/clone="OR\_CBA0057B05"

/tissue\_type="young leaves"

/dev\_stage="2 week old seedlings"

/lab\_host="DH10B T1 phage resistant"

/clone\_lib="OR\_CBA"

/note="Vector: pAGIBAC1; Site\_1: HindIII; Site\_2: HindIII;

drk treated 36 hrs before harvest"

ORIGIN

Query Match 20.0%; Score 494.8; DB 9; Length 706;

Best Local Similarity 86.5%; Pred. No. 4.1e-116;

Matches 607; Conservative 0; Mismatches 7; Indels 88; Gaps 2;

QY 1690 TGTACGATCAAAACCCAGTCCATCCAGCCCTTAAGTACAAGAGGACCTAGGCACCT 1749

DB 1 TGTACGATCAAAACCCAGTCCATCCAGCCCTTAAGTACAAGAGGACCTAGGCACCT 60

QY 1750 CTATGGTGGTCAAAACCGAAGGCTATCTCTTGGTGGAGCCACCATGCAACCCCGAAGA 1809

DB 61 CTATGGTGGTCAAAACCGAAGGCTATCTCTTGGTGGAGCCACCATGCAACCCCGAAGA 120

QY 1810 CTGATATCTGCAATTCAAAGTCTGTCGCTGCCAAGAAACTGAAGAAATCGGCACCT 1869

DB 121 CTGATATCTGCAATTCAAAGTCTGTCGCTGCCAAGAAACTGAAGAAATCGGCACCT 180

QY 1870 TGTCCCTAGTA----- 1881

DB 181 TGTCCCTAGTAAGCCCTACTAGTATCATGTGTCGATATATTTTCTCTTTATTT 240

QY 1882 -----GTAGTAGAGGTTGGACATTC 1902

Db 241 TCACCTTGAACATATGCTTAACCAACCAACATATCAGGTAGAGGTTTGGACATG 300  
Qy 1903 CCGGATGCTTATCAAGAGTTGTCTTCAATGCGAGTACTCTAGTGAAGAGACAC 1962  
Db 301 CCGGATGCTTATCAAGAGTTGTCTTCAATGCGAGTACTCTAGTGAAGAGACAC 360  
Qy 1963 CTCTGTAACCTTTCTGCTAGATCACACAGGAAACAGTGTCTCTGCGGCTGTGCGCC 2022  
Db 361 CTCTGTAACCTTTCTGCTAGATCACACAGGAAACAGTGTCTCTGCGGCTGTGCGCC 420  
Qy 2023 CTATCACCATAACACTGAGGATGATGAGAACAGGACTCCGAAAGACATTTACAGCACTGA 2082  
Db 421 CTATCACCATAACACTGAGGATGATGAGAACAGGACTCCGAAAGACATTTACAGCACTGA 480  
Qy 2083 ATCCCAAGACTCCGATGACTGTTCAGGCTTCAATGCGAGTGAAGTACTCTCTCTGG 2142  
Db 481 ATCCCAAGACTCCGATGACTGTTCAGGCTTCAATGCGAGTGAAGTACTCTCTCTGG 540  
Qy 2143 CCAACAAAGTTTTCAGCACTCCAGTTTCCCTTTTACGACAGGAGGAGGTAACATGCG 2202  
Db 541 CCAACAAAGTTTTCAGCACTCCAGTTTCCCTTTTACGACAGGAGGAGGTAACATGCG 600  
Qy 2203 AGGAGGACATCGAC-TACTCTCTTTGAAGAAAGCGGCTCGCCATCTATCTGCGCAGGCAA 2261  
Db 601 AGGAGGACATCGACTTACTCTTTGAAGAAAGCGGCTCGCCATCTATCTGCGCAGGCAA 660  
Qy 2262 ATGGTTTAACTGTCATCAATTTATGTAGTGTAGTTGAAATCT 2303  
Db 661 ATGGTTTAACTGTCATCAATTTATGTAGTGTAGTTGAAATCT 702

RESULT 5  
LOCUS CA498690  
DEFINITION WHE3246\_E07\_1142T Wheat meiotic anther cDNA library Triticum  
aestivum cDNA clone WHE3246\_E07\_114, mRNA sequence.  
ACCESSION CA498690  
VERSION CA498690.1 GI:24989650  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 646)  
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R.,  
Pham, J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.  
The structure and function of the expressed portion of the wheat  
genomes - Meiotic anther cDNA library  
Unpublished (2002)  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderan@pw.usda.gov  
Sequences have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: T7 primer.

FEATURES  
1..646  
Location/Qualifiers  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE3246\_E07\_114"  
/tissue\_type="Anther"  
/dev\_stage="Meiotic stages pre-meiosis-metaphase I"  
/lab\_host="E. coli DH10B"  
/clone\_lib="Wheat meiotic anther cDNA library"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; Plants

were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN  
Query Match 18.8%; Score 463; DB 6; Length 646;  
Best Local Similarity 83.9%; Pred. No. 6.4e-108;  
Matches 536; Conservative 0; Mismatches 100; Indels 3; Gaps 1;  
Qy 978 ATGATGATCCACTTGAAGAGCAGCAGATGTTTCAGATATAACATCAATATTGCTGCT 1037  
Db 11 ATGATGATCCCAATTTGAAGAGCAGCAGATGTTTCAGATATAACATCAATATTGCTGCT 70  
Qy 1038 TCAGAACAAAGAGATAAAGTGAACCAACACATCTCTCCACAGATTTCTCGAATATTGTCGAA 1097  
Db 71 TCGAGGATGAATTAACAGAACCCACACCTCTGATCGATTCCTCAGCTATGTGAA 130  
Qy 1098 TCTGAGGTGTTAAGGCTTGAACAACTGAAGCAAGTGAAGATGAAGATCTTGTGTTTAAAA 1157  
Db 131 GCTGAGGTTTTAAGGCTTGAACAACTGAAGCAAGTGAAGATGAAGATCTTGTGTTTAAAA 190  
Qy 1158 AAGAAAGCAGAACTAGAGAGCAGTGAAGAGCTGCTCATCTTGTGCGAGGAAGGTTAT 1217  
Db 191 AAGAAATCGGAGCTGAGAGGACACAGAGAGCTGCGCATCTAATTTGGCGAGGAAGGATAT 250  
Qy 1218 GCAGAGGAGTTTATGATGTTGAAGCTTATTAAGAGCTGGAGCTTATGATCCCTCAGTAGTACTT 1277  
Db 251 TCAGATGAATTTAATGATGAGGCTTATGAGTGGGAGCTATTGATCCTGCAATGCTGCTG 310  
Qy 1278 GAACAAATGAGCTCAGTTCAGTTCAGAGTGAAGAGGAAGCTTTAGCCGGAAGATATT 1337  
Db 311 GAACAAATGAGGCTCAGTTCAGTTCAGAGTGAAGAGGAAGCTTTAGCCGGAAGATATT 370  
Qy 1338 CTTGAGAAAGTTGAAGAGTGCAGAAATGCTTTGTAAGAGGAAGCTTGGCTGGAAGATTAC 1397  
Db 371 CTTGAGAAAGTTGAAGAGTTCGAAATGCAATGAGAGGAAGCTTGGCTGGAAGATTAC 430  
Qy 1398 AACAAAGATGATATCGTTTACAATGCTGGAGGGGAGCAGATCTTAACACTAAAGAGGCT 1457  
Db 431 AACAAAGATGATATCGTTTACAATGCTGGAGGGGAGCAGATCTGACGCTCAAGAGGCT 490  
Qy 1458 GAAAGGCTGTTACTTTGGTTCACAGATTCCTGGAATGGTAGATGTTTGGAGAACAAA 1517  
Db 491 GAGAAAGCTCGCATTTTGGTTTAAACAGATTCGCGGAATGGTAGATGTTTGGACACAAA 550  
Qy 1518 ATTGCTCATGGAAAAATGAACGAGGAAGAGGATTTACATATGATGTTGTTAGGCTT 1577  
Db 551 ATCATTCCTTGGGAGATGAAGAGGAAGAGAG---TTCATATATGATGCGTAGGCTT 607  
Qy 1578 TCGTCAATGCTTGAATGATATATGTTTCGTTTCGTCAGGAG 1616  
Db 608 CTCTCAATGCTGGATGAATACATGATTTGTTTCGCGCAGGAG 646

RESULT 6  
LOCUS CA031067  
DEFINITION HX08017r HX Hordeum vulgare subsp. vulgare cDNA clone HX08017  
5-PRIME, mRNA sequence.  
ACCESSION CA031067  
VERSION CA031067.1 GI:24327098  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare





```

5', mRNA sequence.
CA298359
CA298359.1 GI:36068563
EST.
Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 667)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 035 row: E column: 05
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .667
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRPL803505"
/lab_host="DH10B"
/clone_lib="FL8"
/notes="Organ: Developing inflorescence and rachis
(10cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
An unidirectional cDNA library generated from [Developing
inflorescence and rachis (10cm-long)]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

FEATURES             source
source
ORIGIN
Query Match          16.7%; Score 411.8; DB 6; Length 667;
Best Local Similarity 77.2%; Pred. No. 1e-94;
Matches 515; Conservative 0; Mismatches 147; Indels 5; Gaps 1;

QY 56 CTCCTCCCTCTCTCCGCTCTCTCGCATCTGAGGCTCCGATCCGCGCGACCCAGCCAG 115
DB 1 CACTCTCTCGCCCTCTGAGGCGCTGATCGCAAACTCTCCCGCGACAGTCCGCGC 60

QY 116 AATCCGCGCCCGCTCTCGCCCTCCGCTCGAGAGACCGCGCGCGCGAGAGGC 175
DB 61 CCGTCCCGCCCGCGCGCGCGCGCGCTCGAGAGGCTGCGCGCGGTGTGAAGAGG 120

QY 176 -----CTAGTGTCTTTCGCACTTCGCGATGAGTAGCGGTGAAGGACCACTTCACCCAG 230
DB 121 TGTGCTGCGGTGTTCCTCCCTCTCGCTATGAGTAGCGCTGGAAGGACCACTGCAGCAG 180

QY 231 ATGTGACGACATGCGATTCTGCTTACTGGAGCTCAATGTGATTTGGGATGAGTCCGT 290
DB 181 ATGTGACGACGCTGCGATTCTGCTTCTGAGCTCAACGCTGATTTGGGATGAGTCCGT 240

QY 291 GAGCCGACACGACGAGGACAGATCTGCTGAGCTCCGAGGAGTGCCTGGAGGTC 350
DB 241 GAGCCGACACATGTGTGAGGACCGGATGCTGCTGAGCTTCGAGCAGGAGTGCCTCGAGTC 300

QY 351 TACAGGCGGAGGTTCGACAGGCGAAACCGGAGCGCGCCAGCTGCGGAAAGGCCATGCC 410

```

```

Db 301 TACAGGAGGAAGGTTCGACAGCCCAATCGGTGAGGGGCCAGTTTCGGCAGGCCATTCGG 360
QY 411 GAGGCGGAGGAGAGCTCGCGGCATCTGCTCAGCATGCGGCGAGCCCGCCGCTGCAGCTT 470
Db 361 GAGGCGAGGCGGAACTCGCTGGCATCTGCTCGGCCATCGGCGAGCCACCGATACATGTT 420
QY 471 AGACAGTCAATCAGAACTTCATGGCTTAAGAGGAGTTCGAATGCAATTCGTTCCGTAT 530
Db 421 AGACAGTCAATCAGAACTTCATGGCTTAAGGAGGAAATTCGAATGCGGATTCGCCATAC 480
QY 531 TTGGAAGAAATCAAAAAGGTCGAACGATGAAACCAAGTTCGTTGTTTCATGCTCATAGAG 590
Db 481 TTGGAAGAGATGAGAAAGAAAGAACTGAAGATGCGCAATTTGTTGATGTTATAGAG 540
QY 591 CAGATTAAAGAAAATTTTCGTCGTAATAAGGCCAGCGGATTTTGTTCCTTTAAAGTTCG 650
Db 541 CAAATTAGAGGTTGCACTCTGAATCAGGCTGAGGCTTTGCACTTTGAGGATTCCT 600
QY 651 GTTGATCAGTCTGACCTGTCATTAAGAAGCTTGATGAGTTGACGAGGACCTGGGATCC 710
Db 601 GTGGATCAGTCTGATCTGTCTTGTAGAAAGCTGGAGGAGCTAATAAAGAGCTACAAATCC 660
QY 711 CTTTCAGA 717
Db 661 CTTTCAGA 667

```

```

RESULT 10
BG487823/c
LOCUS
DEFINITION
  FM1_73_H06.g1_A003 Floral-Induced Meristem 1 (FM1) Sorghum
  propinquum cDNA, mRNA sequence.
ACCESSION
  BG487823
VERSION
  BG487823.1 GI:13469058
KEYWORDS
  EST.
SOURCE
  Sorghum propinquum
  Sorghum propinquum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
  1 (bases 1 to 617)
  Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M., Marsala,C. and
  Pratt,L.H.
  An EST database from Sorghum: floral-induced meristems
  Unpublished (2000)
  Contact: Cordonnier-Pratt MM
  Laboratory for Genomics and Bioinformatics
  The University of Georgia, Department of Plant Biology
  Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
  Tel: 706 542 1860
  Fax: 706 583 0210
  Email: mmpratt@uga.edu
  Sequences have been trimmed to exclude PolyA, vector and regions
  below Phred quality 16. The threshold for highest quality sequence
  is 20.
  Seq primer: T7
  High quality sequence start: 14
  High quality sequence stop: 613
  POLA=No.
  Location/Qualifiers
  1. .617
  /organism="Sorghum propinquum"
  /mol_type="mRNA"
  /db_xref="taxon:132711"
  /clone_lib="Floral-Induced Meristem 1 (FM1)"
  /note="Organ: Floral-Induced meristems; Vector:
  pluscript II from Lambda Zap II; Site 1: XhoI; Site 2:
  EcoRI; mature plants were placed in a growth chamber for
  15 days with 16 hr darkness and 8 hr light (flowering is
  induced by short-day conditions); 16 days after being
  returned to the greenhouse under natural long days during
  late April/early May, meristems were harvested the
  library was made from poly-A RNA in the cloning vector

```

lambda ZAP II. Clones to be sequenced were prepared by mass excision."

ORIGIN	Query Match	Best Local Similarity	Score	410.4; DB 4; Length 617;	Matches 492; Conservative 0; Mismatches 111; Indels 5; Gaps 1;
QY	101	GGCGACCCAGCAGCAATCGCGCGCCGCTCTCGCCCTCCCGCTCGCAGCAGACGCGCC	160		
Db	611	GGCCCCCCCCCGCGCGCTCCCGCGCCCGCCAGCCGCTCGCAGCAGCAGCGCC	552		
QY	161	GAGCGCGAAGAGGC-----CTAGTGTCTTCGCACTCGCATGAGTAGCGCGTGAAG	215		
Db	551	GGTGGGAAGAGGTGTCTGCGGTGTCCTCCCTCTCGTATAGTAGCGCGTGAAG	492		
QY	216	GACCACTTACAGATGTCAGCAGATCGCATTCGTTCTACTCGAGCTCAATGTATT	275		
Db	491	GACCACTTACAGATGTCAGCAGATCGCATTCGTTCTACTCGAGCTCAATGTATT	432		
QY	276	TGGATGAGGTGCTGAGCCGACACGAGGAGCAGATGCTCTGAGCTCGAGCAG	335		
Db	431	TGGATGAGGTGCTGAGCCGACACATGCGAGGACCGGATGCTCTGAGCTCGAGCAG	372		
QY	336	GAGTGCCTGGAGGTCTACAGCGGAGAGGTGCAACGAGGCAACCGGAGCGCGCCAGCTG	395		
Db	371	GAGTGCCTGGAGGTCTACAGGAGGAGGTGCAACGAGGCAACCGGAGCGCGCCAGCTG	312		
QY	396	CGAAGGCCATCGCGAGGCGAGGAGCTCGCGCGCATCTGCTAGCCATGCGGCGAG	455		
Db	311	CGCGAGGCCATTCGCGAGGCGAGGCTGAACCTCGCTGCGCATCTGCTCGCCATCGCGCAG	252		
QY	456	CGCGCGCTGACGTTAGCAGTCAATCAGAGCTTCATGGCTTAAGAGGAGTTGAAT	515		
Db	251	CCACCGATACATGTTAGCAGTCAATCAGAGTTCATGGTTTACGAGAGGAATGAAT	192		
QY	516	GCAATTGTTCCGTATTGGAAGAAATGAAAGAAAGAGTTCGAACTGAAACCACTTT	575		
Db	191	GCATTTGTTCCCATCTTGGAGGAGATGAGAAAGAAAGTTCGAACTGAAACCACTTT	132		
QY	576	GTTCTATGTCATAGCAGATTAAGAAATTTGCTCTGAAATTAAGCCGCGCATTTGTT	635		
Db	131	CTTGATGTCATAGCAAAATTAAGAAAGTTGTCATCTGAATCAGCGCTCGAGATTTGTA	72		
QY	636	CCCTTTAAAGTTCCGCTGATCAGCTGACCTGTCATTAAGAAAGCTTGATGATTGACG	695		
Db	71	CCCTTTAAAGTTCCGCTGATCAGCTGATCTGTCATTAAGAAAGCTTGATGATTGACG	12		
QY	696	AAGGACCT 703			
Db	11	AAAGAGCT 4			

RESULT 11  
CK193855  
LOCUS  
DEFINITION  
FGA5002274 Triticum aestivum FGAS: Library 3 Gate 6 Triticum  
aestivum cDNA, mRNA sequence.  
ACCESSION  
CK193855  
VERSION  
CK193855.1 GI:39556245  
KEYWORDS  
EST.  
SOURCE  
Triticum aestivum (bread wheat)  
ORGANISM  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 807)  
REFERENCE  
Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,  
Geneswein, B., Graf, R., Gulick, P., Hrycan, L.D., Larocque, A.,  
Links, M.G., McCarthy, E.L., Monro, A., Muzak, I., Nilsson, D.,  
Penniket, C., Roach, J.L. and Sarhan, F.  
Functional Genomics of Abiotic Stress In Wheat and Canola Crops  
Unpublished (2003)

## COMMENT

Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_esta@cs.usask.ca  
This sequence is the direct result of the Base calling software  
Phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [103-775].  
Plate: L3C102 row: A column: 05.

## FEATURES

source

1..807  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: Library 3 Gate 6"  
/note="Organ: Root; Vector: pCMV.SPORT6; Root tissue from  
Saskatchewan, 7 mRNA populations were combined before  
constructing the library; 7 day non-acclimated roots, 1,  
23, and 53 days cold-acclimated at 4C, and 30 minutes, 3  
hours and 6 hours treated roots with 200mM NaCl.  
Non-acclimated and cold-acclimated plants were grown in  
vermiculite while salt stressed plant were grown  
hydroponically. First strand synthesis in this library was  
done in the presence of methylated dCTP thereby protecting  
from internal cleavage with NotI."

## ORIGIN

Query Match	Best Local Similarity	Score	404.2; DB 7; Length 807;	Matches 536; Conservative 0; Mismatches 133; Indels 12; Gaps 4;
QY	1599	ATGTTCTGCTTCAGGAGAAAGAGCAAGAGAGAGCAAAAGGGATCAGAAAGAGCTC	1658	
Db	115	ATGATTGTTGCCAGGAGAAAGCAAGAGAGAGCAAAAGGGATCAGAAAGAGCTC	174	
QY	1659	CAGGATCAGCTCAAAAGCGGAGCAGGAGCTTTGTACCGATCAAAACCCAGTCCAAAG	1718	
Db	175	CACGATCAACTCAAAAGCTGAGCAGGAGCTCTACGGTTCAAAGCCAAAGTCCAAAG	234	
QY	1719	CCCTTAAGTCAAAAGAGGACCTAGCAGCTCTATGGTGGTGCAGAAACCCAGAGCTATCT	1778	
Db	235	CCTCAAGACATAAAGAGAGCGCCCTAGACACTCGATGGTGGTGTAAACCCAGGGTGTCC	294	
QY	1779	CTTGGTGGAGCCACCATGCAACCCCGAAGACTGATATCTGCAATTCAAAGTCTGTTTCT	1838	
Db	295	ATGGTGGAGCCAGATGCAAGCACCACCAACAGACATCTGCATTCAGAAATGTTTCT	354	
QY	1839	GCTGCCAAGAAAACCTGAAGAAATCGGCACCTTTGTCCCTCTAGTAGTAGAGGTTTGGAC	1898	
Db	355	GCTGCCAAGAGAACTGAAGACATTTGCCCAITTTGTCCCTCTAGTAGTAGAGGTTTGGAC	411	
QY	1899	ATTGCCGAGTTGCTATCAAGAGTTGTTCTTCAATGCCAGTACTCTACGTGAGACGGAG	1958	
Db	412	ATTGCCGATCTTCCCATCAAGAGCTATCTTTCAATGCCACTGCTCTTGGTGAACCGAG	471	
QY	1959	ACACCTCGTAAACCTTTTGTCTCAGATCA---CACCCAGGAAACAGTGTCTGTCCAGCGCT	2015	
Db	472	ACACCTCGCAGAGCTTTTGGCAGATCATCGCGCGCCAGCAGTGTCTCCCTCCTCCTCT	531	
QY	2016	GTGGCGCTTATCAACCAATAACACTGAGGATGATGA---GAACAGGACTCCGAGACATTT	2072	
Db	532	GCGCGCTCTGTCAACCAACGACACCGGAGGAGAGAACCGGACCCCAACCCCAAGACATTT	591	
QY	2073	---ACACACTGTAATCCCAAGACTCCGATGACTGTTTACGGCTCCCAATGAGTAGTGC	2129	
Db	592	GGAGCAGCACTGAACCTGAAAGACCCCGATACCGGTGGCGCTCGATGCGAGTGGCAATG	651	



QY 2130 ACTCCCTCTCTGCGCAACAGGTTTCAGCAATCCAGTTCCTGTTTACGACAGCCA 2189  
 DB 652 AGCCAGCTGTGGCTAACAGGTACAGGCTGCTCTGCTCTGCTTACGAGAACCA 711  
 QY 2190 GAGGTAACTTCAGAGGACATCGACTACTCTCTTGAAGAAGCGCGTCCCATCTAT 2249  
 DB 712 GAGCCAACATTCGCGAGCGCATCGAGTATTCCTTCGAAGAAGCGCGTGGCCGCTAT 771  
 QY 2250 CTGGCCAGGCAATGCTTAA 2270  
 DB 772 TTGTCAAGAGAAGTGTGTTAA 792

RESULT 12  
 BG604893 548 bp mRNA linear EST 16-APR-2001  
 LOCUS WHE0944\_F03 L06Zs wheat 5-15 DAP spike cDNA library Triticum  
 DEFINITION aestivum cDNA clone WHE0944\_F03\_L06, mRNA sequence.

ACCESSION BG604893  
 VERSION BG604893.1 GI:13634896  
 KEYWORDS EST  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 548)  
 Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.  
 The structure and function of the expressed portion of the wheat genomes - 5-15 DAP spike cDNA library  
 Unpublished (2000)

JOURNAL Contact: Olin Anderson  
 COMMENT US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105959773  
 Fax: 5105959818

Email: andersow@w.usda.gov  
 Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
 Seq primer: Stratagene SK primer.

FEATURES  
 source Location/Qualifiers  
 1..548  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE0944\_F03\_L06"  
 /tissue\_type="Spike"  
 /dev\_stage="Adult plant"  
 /lab\_host="E. coli SOUR"  
 /clone\_lib="Wheat 5-15 DAP spike cDNA library"  
 /notes="Vector: Lambda Uni-ZAP XR, excised phagemid; Site1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Spikes at 5, 10 and 15 DAP were harvested, total RNA and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."  
 ORIGIN

Query Match 16.3%; Score 402.4; DB 4; Length 548;  
 Best Local Similarity 83.4%; Pred. No. 2.6e-92;  
 Matches 457; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 737 AAAGCAAGTGATAGAAACATTGTAATCTTTGCAATTCCTTATGTGAGGTGCTTGGCATAGA 796  
 DB 1 AAAGCAAGTGATGATCATCTTTGAACACATTACATTCCTTATGTGAGGTGCTTGGCATAGA 60

QY 797 TTTCAAGCAACAGTATATGAGGTGCACCCCTAGCTTGGACGAAGCTGAAGGATCAAGAA 856  
 DB 61 CTTCAACCAACAGTGCATGAGGTGCACCCCTAGCTGAGCAAGCTGAAGGATCAAGAA 120  
 QY 857 CTTGAGCAACACTACCAATTTGAGAGGCTTGTGCTGCCCAACACACACTGCGTGAATGAA 916  
 DB 121 CTTGAGCAACACTACCAATTTGAGAGGCTTGTGCTGCCCAACACACTGCGTGAATGAA 180  
 QY 917 GATCCAAAGGATGCAAAAGCTTTCAAGATTTTGTCTTCTAGCATGCTCGAGCTATGGAATCT 976  
 DB 181 AATTCAGAGGATGCAAAAGCTTTCAAGATTTTGTCTTCTAGCATGCTCGAGCTATGGAATCT 240  
 QY 977 CATGATATCTCCACTTGAAGAGCAGACATGTTTCAGATATTAACATGCAATATTTGCTGC 1036  
 DB 241 TATGATATACCAATTTGAAGAGCAGACATGTTTCAGATATTAACATGCAATATTTGCTGC 300  
 QY 1037 TTCGAAACAAGATTAACAGCAACCACTCTCCACAGATTTTCCCTGGAATATGTCGA 1096  
 DB 301 CTCGGAGCATGAAATTAACGGAACCCCAACCCCTGTGATCGACTTCTCGCTATGTGGA 360  
 QY 1097 ATCTGAGGTGTTAAGGCTTGAACAACTGAAGCAAGTAAGATGAAGATCTTTGTTTAA 1156  
 DB 361 AGCTGAGGTTTAAAGCTTTGACCATTTGAAGGAGCAAAATGAAGGACCTTTGTTTGA 420  
 QY 1157 AAAGAAAGCAGAACTAGAGAGCATAGAGAGCTGCTCATCTTTGCGGAGAGAGTTA 1216  
 DB 421 AAAGAAATCAGAACTGGAGAGCAGACAGACCGCGCATCTAAATTTGGCGAGAGGATA 480  
 QY 1217 TCAGAGAGATTTAGCATTTGAAGCTTATGAAGCTGAGGCTATTGATCCCTACTAGTACT 1276  
 DB 481 TTCAGATGAATTTAAACATTTGAGGCTATTGAGTCGGAGCTATTGATCTCGCATTTGCT 540  
 QY 1277 TGAACAAA 1284  
 DB 541 GGAACAAA 548

RESULT 13  
 CNS0A2G0  
 LOCUS

DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTPGH202C09 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION BX827777  
 VERSION BX827777.1 GI:42459968  
 KEYWORDS HTC; GSUT\_cDNA  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 1986)

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 Unpublished  
 2 (bases 1 to 1986)

JOURNAL Direct Submission  
 REFERENCE Genoscope.  
 AUTHORS Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 JOURNAL - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)  
<http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis>

# FEATURES

source

Location/Qualifiers  
 1..1986  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTPGH202C09"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1986)  
 /gene="At4g26760"

gene

## ORIGIN

Query Match 16.1%; Score 398.4; DB 3; Length 1986;  
 Best Local Similarity 56.0%; Pred. No. 46-91;  
 Matches 841; Conservative 0; Mismatches 646; Indels 15; Gaps 4;  
 240 ACATCGCATTCCTCTACTGAGCTCAATGTGATTTGGGATGAGTGGTGAGCCGAC 299  
 139 ACTTGTGGCACCTTACTACAGAAATTGAGGAAATCTGGATGAGTTGGGGAGATGAT 198  
 300 ACAGCAGGACAGATGCTGTGAGCTCGAGCAGGAGTGTGAGGAGTCTACAGCGG 359  
 199 GAAGAACAGACAACTGCTTACAGATTGAGGAGAGTGTCTGAATGTTTACAAAAG 258  
 360 AAGTTCGACGACGCGAACCGGAGCGCGCCAGCTCGGAGGCCATCGCGAGGCGAG 419  
 259 AAGTTGAGCTAGCGCGAAATCTCTGCGAGAGTCTTTCAGACTTTGTGATGCCACT 318  
 420 CGAGAGCTCGCGGATCTGTCTAGCCATGAGCGCGCCGCTGACGCTTAGACAGTCA 479  
 319 GTTGAATTTTCAATCTCACACTGCTCTCGGGGAAAGAGCTATATCGACATTC--CT 375  
 480 ATCAGAGCTTCATGGCTTAAGAGAGGAGTTGAATGCAATTTGTCGATTTTGAAGAA 539  
 376 GATTAAGCTTCAGGACGATCAAGAACAACTTTCTGCAATAGCACCTGCGCTTGAACA 435  
 540 ATGAAAAAGAAAAGTCAACGATGGAACAGTTTGTTCATGTATAGAGCAGATTAAG 599  
 436 CTTTGGCAACAGAGAGAAAGGTCGGGCAATTTCTGATGTACATCATCAGATTGAG 495  
 600 AAAATTTCTGTGAATTAAGCCAGCCGATTTGTTTCCCTTTAAAGTTCCGGTTGATCAG 659  
 496 AAGATATGTTAGAGAT---CGCTGGCGATTGAAACAAATGGACCTCATGTGTTGATGAG 552  
 660 TCTGACCTGTCAATTAAGAAAGCTTATGATGATGAGAGGACCTGGATCCCTTCAGAG 719  
 553 ACTGACTTGTCTTAAAGAGATTAGATGACTTCCAGAGGAACTCCAAGAGCTCCAGAAA 612  
 720 GAGAGAGCGATCGGCTAAGCAAGTGTAGAACTTTGAATTTCTTCATTTCTTATGT 779  
 613 GAGAGAGTGTAGTTTACAGAGGTTGCTCGAGTTTGTAGACGCTCATGATCTATGT 672  
 780 GAGTGTCTTGGATAGATTTCAGCAAAACAGPATATGAGGTGCACCCCTAGCTTGGACGAA 839  
 673 TCCGCTCTTGGTTAGATTCTTAAAGCACTGTCCAGGAAAGTTTATCCGAGCTTAGATGAG 732  
 840 GCTGAAGG-----ATCAAGNACTGAGCACTACAAATTTGAGAGCTTGTCTGCTGCC 893  
 733 GAAATGTTGTTTCAACCAAGAGCATTTAGCAATGAAACACTTGCAGGTGGCTAAGACT 792  
 894 GCAACACAGATCGCTGAATGAAGTCAAAAGGATGCAAAAGCTTCAAGATTTTGTCTCT 953  
 793 GTCTTGACTTTAAGAGAGATTAATGCAAGACTTTAAGAGCTTCAAGACTAGTACT 852  
 954 AGCATGCTCGAGCTATGGAATCTCATGATCTCACTTTGAGAGCAGCAGATGTTTCAAG 1013

Db 853 CAGCTAACTGATCTATGGAATCTGATGATCTTCTTGATGAGGAAAGGGAGCTTTTTCAC 912  
 Qy 1014 AATATAACATCATATATTTGCTGCTTTCAGAAACAGAGATAAATGACCAACACCCCTCTCC 1073  
 Db 913 CATGTACCTCTAAATTTTCAGCTTTCAGTTTCATGAAGTACTGCTCTCTGGTCTCTCGCA 972  
 Qy 1074 ACAGATTTTCTTGAATTTATGTCGAATCTGAGGTGTTAAGGCTTGAACAACTGAAAGCAAGT 1133  
 Db 973 CTTGATCTTATCGAACAGGCTGAGGTGGAAGTGTGATGAGCTTGTATCACTGAAATCTAGC 1032  
 Qy 1134 AAGATGAAGATCTTGTGTTTAAAAAAGAAACAGAACTAGAGAGCATAGAGAGCTGCT 1193  
 Db 1033 AGAATGAAGGAAATTCATTCAAAAGCAATCAGAGCTTTGAGGAGATATATGCTGCTGCT 1092  
 Qy 1194 CATCTTGTGGCAGGAGGTTATGAGAGAGTTTACAGATTTAGCTTGAAGCTTATGAGCTGGA 1253  
 Db 1093 CACATAGAAATPAAAAACCAGAGGTTGTTGCTGAGAGGATCATGTCTGTTGATTTGCTGGG 1152  
 Qy 1254 GCTATTGATCCTCACTAGTACTTTGAACAAATTTGAAGCTCACATTTGCAACAGTGAAGAG 1313  
 Db 1153 AATCTGAACTACTGAGTTACTGCTGACATGATAGCCAGATCGCAAGGCCAGAG 1212  
 Qy 1314 GAAGCTTTTACCCGGAAGGATTAATCTTGAAGAACTTGAAGATGCGCAAAATGCTTTGTGAA 1373  
 Db 1213 GAAGCTTTTGTAGTAAAGAAATTTTAGACCGTGTAGAGAAATGATGTCAGCTTGTGAG 1272  
 Qy 1374 GAGGAAGCTGGCTGGAGGATTACAAAGAGATTAATCGTTTCAATGCTGCGAGGGGA 1433  
 Db 1273 GAAGAGAGCTGGCTCGAAGATTACAAATCGAGATCAGAAACAGGTATAGTGCAGCGAGGT 1332  
 Qy 1434 GCACATCTAACACATAAGAGGCTGAAAGGCTCGTACTTTGGTCAACAAAGTTCCTTGA 1493  
 Db 1333 GCGCATTTGATCTCAAGAGAGCTGAGAAAGCTCGGATTTGCTGTCAGCAAGTACTGCT 1392  
 Qy 1494 ATGTGATGATTTTGAAGCAAAATTTGCTGATGAGAAATTAAGCAAGGAGGAGGAT 1553  
 Db 1393 ATGTTTGCACGTTGATTTGCTTAAACCCGGGATGGAAGAGGAGAACAGTA---TGTCC 1449  
 Qy 1554 TTCACATATGATGCTGTTAGCTTTTCGTCATGCTTGTGATGATATATATGCTGCTGTCAG 1613  
 Db 1450 TTTGAGTACGATGAGTTCCTTTACTCGCCATGCTAGTGAATACATATGCTTTAGGCA 1509  
 Qy 1614 GAGAAAGAGCAAGAGAGAGAGAGCAAAAGGATCAGAAAGCTCCAGGATCAGCTCAAA 1673  
 Db 1510 GAACGAGAGATGAGAGCGAGGCTGAAGAAACAGAAAGAGCAGCAAGACCCACAC 1569  
 Qy 1674 GCGAGAGAGGAGCTTTGTAGCGATCAAAACCCAGTCCATCCAGGCCCTTAAGTACAAAG 1733  
 Db 1570 ACAGATCAAGAAATCCGCTTTGGTCTTAAACCAAGCCCTGCAAGACCCGTTAGTGCACAG 1629  
 Qy 1734 AA 1735  
 Db 1630 AA 1631

## RESULT 14

CNS09YQA

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSLTUS592C11 of Adult vegetative tissue of strain col-0 of

Arabidopsis thaliana (thale cress).

Accession

Version

Keywords

Source

Organism

Reference

Authors

CNS09YQA 2195 bp mRNA linear HTC 06-PEB-2004  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSLTUS592C11 of Adult vegetative tissue of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX831110  
 GI:42454972  
 HTC; GSLT cDNA  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1 (bases 1 to 2195)  
 Castellani,V., Aury,O.M., Jaillon,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
JOURNAL 2 (bases 1 to 2195)  
REFERENCE 2 (bases 1 to 2195)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequences :  
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source

Location/Qualifiers  
1..2195  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTSL592C11"  
/tissue\_type="Adult vegetative tissue"  
/plasmid="pCMVSPORT\_6"  
1..2195  
/gene="At5g5230"

## gene

## ORIGIN

Query Match 16.0%; Score 395.6; DB 3; Length 2195;  
Best Local Similarity 55.7%; Pred. No. 2.2e-90;  
Matches 846; Conservative 0; Mismatches 659; Indels 15; Gaps 4;  
QY 222 CTTACACAGATGCGACGATGCGATTCGTTCTACTGGAGCTCAATGTGATTTGGGAT 281  
DB 146 CCTCATCTTTGGGAAATTACTTGTGTACTTACTTTGAGAAGTTGCAGGAAATCTGGAT 205  
QY 282 GAGTTCGTTAGCCACACGACGAGGACAGATGCTGCTGGAGCTCGCAGCAGAGTGC 341  
DB 206 GAAGTTGGTGGAGTGTATGATGAACGAGACAACTGCTTCTTCAGATAGCAAGAGTGT 265  
QY 342 CTGGAGGTCTACAGCGGAGGTGCGACCGGCGAACCGGAGCGCGCCAGCTGCCGAAG 401  
DB 266 CTTGACGTTTACAGAGAAAGTTCGACGAGGTGCGAAATCCGAGCTGAGCTTCTTCAA 325  
QY 402 GCCATCCCGAGGCGGACGAGCTGCGCGGATCTGCTAGCCATGGCGAGCGCGCC 461  
DB 326 ACCTTGTGAGTCTAATGCTGAATCTTCAGGCTCACAATGCTCTTCGAGACAAAGC 385  
QY 462 GTCCAGCTTAGACAGTCAATCAGAGCTTCATGGCTTAAGAGGAGTGAATCAATT 521  
DB 386 TT---AGTTGGCATTCGGATAGTCTTCAGGAACGATTAAAGAACAACTTGTGCAATA 442  
QY 522 GTTCCGTTATTGGAAGAAATGAAAAGAAAGGTCGACGATGGAACCAAGTTGTTCAT 581  
DB 443 GCACCGGCTTTGAAACAACTGTGGCAACAGAAAGAGAGAGTCCGAGAGTCTCTGAT 502  
QY 582 GTCATAGACAGATTAAAGAAATTTGCTCTGAAATAAGCCAGCCGATTTTGTTCCTTT 641  
DB 503 GTACATCAACAGATTCAGAGATATGAGAGATAT---TGCTGGAGTTTGCAATGAG 559  
QY 642 AAGTTCCGTTGATCAGTCTGACCTGCTCATTTAAGAAAGCTTGAATGAGCAGAGGAC 701  
DB 560 GTTCCCTATAGTCGATGATGATTTGTCTACTGAAGAAATTTAGACGATTTCCAGACCAA 619  
QY 702 CTGGAATCCCTTCAGAGGAGGAGAGCGATCGGCTAAAGCAAGTATAGAACATTTGAAT 761

RESULT 15

DB 620 CTCCAAGAGCTCCAGAGAGAGAGAGTGCAGGCTGCGCAAGGTTTAGAGTTTGTGAGT 679  
QY 762 TCTTTGCAATTCCTTATGTGAGGTGCTTGGCATAGATTTCAAGCAAAAGTATATGAGGTG 821  
DB 680 ACTGTTCAATGATCATGTGCTGCTTCTTGGTTTGAATTTCTTAAGCACCCTCAGGAT 739  
QY 822 CACCTAGCTTTGGAGCAAG-----CTGAAGGATCAAAAGACCTTGAGCAACATCAATT 875  
DB 740 CATCGAGCTTAGATGAAGATACCAAGTGTCCAGTCTTAAGAGCATTAGCAATGAGACTCT 799  
QY 876 GAGAGCTTGTGCTGCTGCGCAAAACAGACTCGGTGAATGAAGATCAAAAGGATGCAAAAG 935  
DB 800 TCAAGTTGGCTAAGACCGCTCTTGACTCTTAAAGATGATNAAGACACAGACTTCAAAAG 859  
QY 936 CTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATATCCACTTGA 995  
DB 860 CTTCAAGAGCTGCTTACTCAGCTAATGACTGTGGAATCTGATGATATCTCTGATGAG 919  
QY 996 GAGCAGCAGATGTTTTCAGATATATAACATGCAATATTGCTGCTTCAGAAACAGAGATACT 1055  
DB 920 GAAAGAGAGCTTTTGTATCATGTTTACCTGTGACATTTTCATCTCAGTCGATGAGTCACT 979  
QY 1056 GAACCAAAACACCTCTCCACAGATTTCTCGAATTTATGTGCAATCTGAGGTGTTAAGCTT 1115  
DB 980 GTGCCAGGTGCTCTTGGCCGCTGATTTGATTTGAGCAGGCTGAGGTGGAAGTTGATAGGCTT 1039  
QY 1116 GNAACATGAAAGCAAGTAAAGTAAAGATCTTGTTTTAAAAAAGAAAGACAGCAACTAGAA 1175  
DB 1040 GACCAAGCTGAAAGCTAGCGGAATGAAGAAATTTGCTTCAAGAAAGCAATCTGAGCTTGA 1099  
QY 1176 GAGCATAGAAGACGTGCTCATCTTGTGGCGAGGAAGTTATGACAGAGGAGTTTACGATT 1235  
DB 1100 GAGATATATGCTCGTGCCCATGTAGAAGTTAACCCGGAATCTGCTGAGAGAAATCATG 1159  
QY 1236 GAAGCTATTGAAGCTGGAGCTATTGATCCCTCAGTACTGTAACCAATTTGAAGCTAC 1295  
DB 1160 TCGCTGATTGATTCTGGAAACGTTGAGCTTACTGAAATTTATTTGGCAGACATGATAGCCAG 1219  
QY 1296 ATTGCAACAGTCAAGAGAGAGAGCTTTTAGCCGGAAGGATATTCTTTGAGAAAGTTGAAGA 1355  
DB 1220 ATATCAAAAGGCTAAGAAAGAGCATTTAGTAGAAAGATATATTGGACCGAGTCGAGAA 1279  
QY 1356 TGGCAAAATGCTTTGTGAAGAGAGAACCTGGCTGGAAGATTACAAACAAAGATCAATCGT 1415  
DB 1280 TGGATGTGAGCTTTGTGAGAGAGAGAGCTGGCTTAGAAGACTACAATCGGGATCAGAACAG 1339  
QY 1416 TACAATGCTGGAGGGGAGACATCTAACATAAGAGGGCTGAAAGGCTGCTACTTTG 1475  
DB 1340 TACAGCGCAAGCAGAGGTGCACATTTGAATCTCAAGAGAGCTGAGAAAGCTCGGATTCG 1399  
QY 1476 GTCACAAAGATTCTGGAATGTTAGTATGTTTGAACAAACAAATTTGCTCATGGAAAAAT 1535  
DB 1400 GTTACAGAGATTCTCGCCATGTTGACATATTAGTTGCCAAGACCGGGCTTTGGGAGAA 1459  
QY 1536 GAAACAGAGAAAGGAGGATTTCAATATGATGTTGTTAGCTTTTCGTAATGCTTGTGAA 1595  
DB 1460 GAAACAGCA---TGTCCTTTGCTTACGATGTTGTTCTCTGCTAGCTATGCTAGACGAG 1516  
QY 1596 TATATGTTGCTGCTCAGAGAAAGAGCAAGAGAGAGAGAGACAAAGGGATCAGAGAGAG 1655  
DB 1517 TACGGTATGCTTAGGCAAGAACGAGAGAGAGAGAACCGGAGGCTGAGGGAAACAAAGAGAG 1576  
QY 1656 CTTCCAGGATCAGCTCAAGCCGAGGAGGAGGCTTTGTACGGATCAAAACCCAGTCCATCC 1715  
DB 1577 GTTCAAGAACACCCACAGTAGAGCAAGAAATCTGCTTTAGCACCGGCAAGCCCTGCA 1636  
QY 1716 AAGCCCTTAAGTACAAAGAA 1735  
DB 1637 AGACCGGTGCTGCTAAGNA 1656



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2004, 18:55:09 ; Search time 207 Seconds  
(without alignments)  
8474.526 Million cell updates/sec

Title: US-10-619-685-1  
Perfect score: 2468  
Sequence: 1 aaaaaattttcaattcac.....aaaaaaaaaaaaaaaaaaaaa 2468

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.4	4.1	267	4	US-09-313-294A-843 Sequence 843, App
2	92	3.7	298	4	US-09-313-294A-950 Sequence 950, App
3	73.8	3.0	7218	1	US-08-232-463-14 Sequence 14, Appl
4	59	2.4	296	4	US-09-313-294A-6877 Sequence 6877, App
5	47.6	1.9	3763	4	US-09-319-039-243 Sequence 243, App
6	47.6	1.9	5185	4	US-09-976-594-640 Sequence 640, App
7	46	1.9	1425	2	US-08-578-551-2 Sequence 2, Appli
8	46	1.9	1425	3	US-09-190-982-2 Sequence 2, Appli
9	46	1.9	1425	3	US-09-408-257-2 Sequence 3, Appli
10	43.4	1.8	16550	4	US-08-916-4218-3 Sequence 3, Appli
11	43.4	1.8	16550	4	US-09-692-570-3 Sequence 3, Appli
12	43.2	1.8	3095	6	5231168-1 Patent No. 5231168
13	43	1.7	7218	1	US-08-232-463-14 Sequence 14, Appl
14	42.2	1.7	1230025	4	US-09-198-452A-61 Sequence 1, Appli
15	42	1.7	2082	4	US-09-818-780-67 Sequence 67, Appl
16	42	1.7	2277	1	US-08-676-967-2 Sequence 2, Appli
17	42	1.7	2277	1	US-08-676-974-2 Sequence 2, Appli
18	42	1.7	2277	2	US-09-098-487-2 Sequence 2, Appli
19	41.8	1.7	798	4	US-09-724-797-69 Sequence 69, Appl
20	41.6	1.7	786	4	US-09-252-991A-6992 Sequence 6992, App
21	41.6	1.7	828	4	US-09-252-991A-3643 Sequence 3643, App
22	41.6	1.7	1458	4	US-09-252-991A-3817 Sequence 3817, App
23	41.6	1.7	1779	4	US-09-252-991A-3864 Sequence 3864, App
24	41.6	1.7	2412	4	US-09-252-991A-7043 Sequence 7043, App
25	41.6	1.7	2472	4	US-09-252-991A-7084 Sequence 7084, App
26	41	1.7	855	4	US-09-252-991A-10200 Sequence 10200, A
27	41	1.7	978	4	US-09-252-991A-14837 Sequence 14837, A

C 28	41	1.7	1158	4	US-09-252-991A-9861	Sequence 9861, Ap
C 29	41	1.7	1230	4	US-09-252-991A-14780	Sequence 14780, A
C 30	40.6	1.6	724	4	US-08-956-171E-832	Sequence 832, App
C 31	40.6	1.6	724	4	US-08-781-986A-832	Sequence 832, App
C 32	40.6	1.6	999	4	US-09-252-991A-14970	Sequence 14970, A
C 33	40.4	1.6	1092	4	US-09-252-991A-9447	Sequence 9447, Ap
C 34	40.4	1.6	1287	4	US-09-252-991A-9384	Sequence 9384, Ap
C 35	40.4	1.6	1785	4	US-09-252-991A-9417	Sequence 9417, Ap
C 36	40.4	1.6	1842	4	US-09-252-991A-9328	Sequence 9328, Ap
C 37	40.2	1.6	417	4	US-09-252-991A-7112	Sequence 7112, Ap
C 38	40.2	1.6	1410	4	US-09-252-991A-9010	Sequence 9010, Ap
C 39	40.2	1.6	1695	4	US-09-252-991A-8785	Sequence 8785, Ap
C 40	40	1.6	1611	4	US-09-600-099-2	Sequence 2, Appli
C 41	40	1.6	6436	4	US-09-621-976-780	Sequence 1, Appli
C 42	39.8	1.6	704	4	US-09-252-991A-15275	Sequence 780, App
C 43	39.8	1.6	1248	4	US-09-252-991A-15326	Sequence 15275, A
C 44	39.8	1.6	2049	4	US-09-252-991A-15326	Sequence 15326, A
C 45	39.8	1.6	2127	4	US-09-252-991A-15511	Sequence 15511, A

ALIGNMENTS

RESULT 1  
US-09-313-294A-843  
; Sequence 843, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 843  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700549927H1  
US-09-313-294A-843

Query Match	4.1%	Score 100.4	DB 4	Length 267
Best Local Similarity	65.2%	Pred. No. 5.7e-18		
Matches 163	Conservative 0	Mismatches 86	Indels 1	Gaps 1
QY	1189	GTGCTCATCTTGTGGCGAGGAGGTTATGCGAGGAGTTTATGACATTTGAAGCTTATTGAAG	1248	
DB	18	GTGCTCATTTGTAATAGACACAGCTGCTCATGAGAAATATTTGGCACTGATTGAGG	77	
QY	1249	CTGAGCTATTGATCCCTCACTAGTACTTGAACAATTTGAAGCTTCACTTGCACAGTGA	1308	
DB	78	CAGGTAACATAGAACCCATCAGAACTAATTGCGAGATATGATGCCAGATAGCAAAAGCAA	137	
QY	1309	AGAGGAAGCTTTAGCCGAGGATATCTTGGAGAACTTGAAGATGCAAAATGCTT	1368	
DB	138	AGGAAGAAGCTTTAGTAAAGATATCTTGTATTAAGTAGAAGATGATGCTTCGCGT	197	
QY	1369	GTGAGAGGAAGCTTGGCTGGAGATTAACAACAAGATGATATCGTTACAAATGCTGGGA	1428	
DB	198	GTGAGAGGAAGCTTGGCTTGAAGATTAACCGGATTAACCGGATGACACAGTATTAAGCTTAGCC	256	
QY	1429	GGGAGCACA	1438	
DB	257	GAGTGCCCA	266	

RESULT 2  
US-09-313-294A-950

; Sequence 950, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PI-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 950  
; LENGTH: 298  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700550124H1  
US-09-313-294A-950

Query Match 3.7%; Score 92; DB 4; Length 298;  
Best Local Similarity 58.3%; Pred. No. 1.3e-15;  
Matches 161; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
QY 902 ACTGCGTGAATGAAGATGCAAGGATGCAAAAGCTTCAAGATTTTCTTAGCATGCT 961  
Db 17 ACTTGAAGAGAAAATGCAAGAGTTTGCAGAGATTCAGCTTAGCTTCCAGCTATC 76  
QY 962 CGAGCTATGGATCTCATGGATCTCCACTTTGAAGAGCAGCAGATGTTTTCAGAAATATAAC 1021  
Db 77 TGATCTTTGGAACTTAATGATTTCTCTGTGAGGAGAACAGCAACCTTTTGGTCAGGTAAC 136  
QY 1022 ATGCAATATGCTGCTTCAGAACAGAGATACTGAACCAACACCCCTCTCCACAGATTT 1081  
Db 137 ATGCATAATGCTCCTCAACATTGGATGATGTAGTATCCAGAGACTCTGGCTCTTGATGT 196  
QY 1082 CCGTAATATGTCGAATCTGAGGTGTTAAGGCTTCAACCACTGAAAGCAAGTAAGATGAA 1141  
Db 197 AATTCAGACGGCGAACTCGAAGTTGAAAGGCTTGAATCAAGAAAGCTAGTAGGATGAA 256  
QY 1142 AGATCTTGTTTTAAAAAAGAAAGCAGAACTAGAAGA 1177  
Db 257 GGATATTGCAATCAAGAAACAGACTGAACTTGAAGA 292

RESULT 3  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 3.0%; Score 73.8; DB 1; Length 7218;  
Best Local Similarity 4.7%; Pred. No. 1.2e-09;  
Matches 18; Conservative 231; Mismatches 138; Indels 0; Gaps 0;  
QY 1269 CTAGTACTTGACAAATTTAGCTCACATTCGAACAGTGAAGAGAGCTTTTAGCCGG 1328  
Db 1466 CAAGTAGTTAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRR 1407  
QY 1329 AAGGATATTTCTTGAGAAAGTTGAAAGATGGCAAAATGCTTGAAGAGAGCAAGCTGGCTG 1388  
Db 1406 RRR 1347  
QY 1389 GAAGATTACAAAGATGATACTGTTACAACTCTGGAGGGAGCACATCTAACACTA 1448  
Db 1346 RRR 1287  
QY 1449 AAGAGGGCTGAAAGGCTCGTACTTTGTCACCAAGATTCCTCGAATGGTAGATGTTTG 1508  
Db 1286 RRR 1227  
QY 1509 AGAAACAAAATTTGCTGCATGCAAAAAATCAACGAGAGAGGAGGATTTTCACATATGATGGT 1568  
Db 1226 RRR 1167  
QY 1569 GTTAGCCCTTTCGTCATGCTGATGAATATATGTTGCTGTCAGGAGAAAGACAAG 1628  
Db 1166 RRR 1107  
QY 1629 AAGAGAGACAAAGGGATCAGAAGAAG 1655  
Db 1106 RRR 1080

RESULT 4  
US-09-313-294A-6877  
; Sequence 6877, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PI-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 6877  
; LENGTH: 296  
; TYPE: DNA  
; ORGANISM: Zea mays



APPLICATION NUMBER: DK 0811/93  
FILING DATE: 06-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 95/02044  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4006.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1425 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Aspergillus aculeatus

Query Match 1.9%; Score 46; DB 2; Length 1425;  
Best Local Similarity 47.3%; Pred. No. 0.022;  
Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 133 CGCCCTCCCGCTCGACGAGACCGCGGAGGCGGAGAGCCCTAGTGTCTTCGACACC 192  
DB 578 CGCCCTGGCTGCTGCGCGCTACCGTGCAGCGGTGACCTTGTCTTGTAGAGCTCGCGC 519

QY 193 TCGCGATGAGTACGCGGTGAGGACGAGTTCACGACATGCGATTCGC 252  
DB 518 TGCCGAGTGCCTGCGCGGTAGAGATGCTCCAGCTGAGCGGACAGCTTCTGGCGG 459

QY 253 TTCTACTGAGCTCAATGTGATTTGGGATGAGTCCGTCGAGCCGACACGAGGAGCA 312  
DB 458 TGCTCGCGGGGTGTAGATGCTGTGCGGTGCGGAGAGGCGCGAGCTCCGAGGAGA 399

QY 313 GGATGCTGTGAGCTCGACGAGAGTGCCTGAGAGGTCTACAGCGGAGGTTCGACACAGG 372  
DB 398 AGACCCAGAGATCGCGGATCCGCTGTCGAAGTCCAGGTTCAAGGTGAGCTGCCACGG 339

QY 373 CGAACCGGAGCGCGCCAGCTCGCGAGGCGCATCGCGAGGCGAGGAGGC 426  
DB 338 TGACCGGGGTCAAGTAGGAGAGCTGCTTGGCTTGGCGGTGGTTCACGACGCTGC 285

RESULT 8  
US-09-190-982-2/c  
Sequence 2, Application US/09190982  
Patent No. 5998190  
GENERAL INFORMATION:  
APPLICANT: Dalboge, Henrik  
APPLICANT: Christgau, Stephan  
APPLICANT: Andersen, Lene N.  
APPLICANT: Kofod, Lene V.  
APPLICANT: Kauppinen, Sakari M.  
APPLICANT: Nielsen, Jack B.  
APPLICANT: Dambmann, Claus  
TITLE OF INVENTION: An Enzyme with Protease Activity  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 59981900 No. 5998190disk of No. 5998190th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/190,982  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/578,551  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 95/02044  
FILING DATE: 19-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4006.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1425 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Aspergillus aculeatus

US-09-190-982-2

Query Match 1.9%; Score 46; DB 2; Length 1425;  
Best Local Similarity 47.3%; Pred. No. 0.022;  
Matches 139; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 133 CGCCCTCCCGCTCGACGAGACCGCGGAGGCGGAGGCGCTAGTGTCTTCGACACC 192  
DB 578 CGCCCTGGCTGCTGCGCGCTACCGTGCAGCGGTGACCTTGTCTTGTAGAGCTCGCGC 519

QY 193 TCGCGATGAGTACGCGGTGAAAGACGAGTTCACAGATGTTCGACACATGCGATTGC 252  
DB 518 TGCCGAGTGCCTGCGCGGTAGGAGATGCTCCAGCTGTAGCGGACAGCTTCTGGCGG 459

QY 253 TTCTACTGAGCTCAATGTGATTTGGGATGAGTCCGTCGAGCCGACACGAGGAGCA 312  
DB 458 TGCTCGCGGGGTGTAGATGCTGTGCGCGGTGCGGAGGCGCGAGCTCCGAGGAGA 399

QY 313 GGATGCTGTGAGCTCGACGAGAGTGCCTGAGAGGTCTACAGCGGAGGTTCGACACAGG 372  
DB 398 AGACCCAGAGATCGCGGATCCGCTGTCGAAGTCCAGGTTCAAGGTGAGCTGCCACGG 339

QY 373 CGAACCGGAGCGCGCCAGCTCGCGAGGCGCATCGCGAGGCGAGGAGGC 426  
DB 338 TGACCGGGGTCAAGTAGGAGAGCTGCTTGGCTTGGCGGTGGTTCACGACGCTGC 285

RESULT 9  
US-09-408-257-2/c  
Sequence 2, Application US/09408257  
Patent No. 6190905  
GENERAL INFORMATION:  
APPLICANT: Dalboge, Henrik  
APPLICANT: Christgau, Stephan  
APPLICANT: Andersen, Lene N.  
APPLICANT: Kofod, Lene V.  
APPLICANT: Kauppinen, Sakari M.  
APPLICANT: Nielsen, Jack B.  
APPLICANT: Dambmann, Claus  
TITLE OF INVENTION: An Enzyme with Protease Activity  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:



RESULT 10  
US-08-916-421B-3/c  
; Sequence 3, Application US/08916421B  
; Patent No. 6503729  
; GENERAL INFORMATION:

## RESULT 11

	Query Match	1.8%	Score 43.4;	DB 4;	Length 16550;
	Best Local Similarity	47.9%;	Pred. No. 0.55;		
	Matches 125;	Conservative	0;	Mismatches 136;	Indels 0; Gaps 0;
Oy	1120	AACGTGAAGCAAGTAAAGATCTTGTTTTAAAAAAGAAGCAGAACTAGAAAGC	1179		
cb	495	NCTCTTTATCTCTTTGTCGTCCTCTTCAGAAAATAAAGAATTAAGATTTCAGAAATG	4236		



[illegible]

```

; LOCATION: (6900001)..(7050000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7050001)..(7200000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7200001)..(7350000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7350001)..(7500000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7500001)..(7650000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7650001)..(7800000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7800001)..(7950000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (7950001)..(8100000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8100001)..(8250000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8250001)..(8400000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8400001)..(8550000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8550001)..(8700000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8700001)..(8850000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (8850001)..(9000000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (9000001)..(9150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature

```

	Query Match	1.7%	Score 42.2;	DB 4;	Length 1230025;
	Best Local Similarity	54.1%;	Pred. No. 18;		
	Matches 86;	Conservative 0;	Mismatches 73;	Indels 0;	Gaps 0;
Qy	1142	AGATCTCTGTTTAAARAAGACGAGACTAGACATAGACGCTGCTCATCTCTG	1201		
Db	348227	AGAAATGCTTCAAAATAACATCCCTTACCCAAATTAAAGACCAGAAAGTGGGGCTCTTTT	348286		
Qy	1202	TGGCGAGGAAGGTTATGCGAGGAGGTTTGTAGCATTGAAGCTATTGAAGCTGGAGCTATTGA	1261		
Db	348287	TTCTCAATTAGATTTTGTATGAGACTTCGAGAAATCTAGCTTAAGAATACGACTCTGTTGA	348346		
Qy	1262	TCCCTCACTAGTACTTGAAACAAATTGAAGCTCAATTGC	1300		
Db	348347	GCCTAAAGTAGAAATTTTCTGAAGGGAAACTTAACATAGC	348385		

```

, PRIOR APPLICATION NUMBER: US 60/192,736
,
, PRIOR FILING DATE: 2000-03-28
,
, NUMBER OF SEQ ID NOS: 98
,
, SOFTWARE: PatentIn version 3.0
,
, SEQ ID NO: 67
,
, LENGTH: 2082
,
, TYPE: DNA
,
, ORGANISM: Thermus thermophilus
US-09-818-780-67

```



Db CCTCTCTCCGCTCTCTCGCATCTGAGCTCCGATCGCGGACCCAGCGAGAATCC 120  
QY GCGCCCGCTCTGCGCTCTCCGCTCGAGAGACCGCGCGAGCGGCGAAGGCGCTAGT 180  
Db GCGCCCGCTCTGCGCTCTCCGCTCGAGAGACCGCGCGAGCGGCGAAGGCGCTAGT 180  
QY GTTCTTCGCACCTCGCGATGAGTAGCGGCTGAAGGACCACTTCACAGATGTCGACGA 240  
Db GTTCTTCGCACCTCGCGATGAGTAGCGGCTGAAGGACCACTTCACAGATGTCGACGA 240  
QY CATCGATTCGCTTCTACTGAGCTCAATGTGATTTGGGATGAGTCCGTCAGCCCGACA 300  
Db CATCGATTCGCTTCTACTGAGCTCAATGTGATTTGGGATGAGTCCGTCAGCCCGACA 300  
QY CGACGAGGACAGGATGCTGTCGAGCTCGAGCAGGAGTGCCTGAGGCTTACAGGCGGA 360  
Db CGACGAGGACAGGATGCTGTCGAGCTCGAGCAGGAGTGCCTGAGGCTTACAGGCGGA 360  
QY AGGTCGACCGAGCGAAGCGAGCGCGCCAGCTCGGAGGCGCATCCGAGGCGGAGG 420  
Db AGGTCGACCGAGCGAAGCGAGCGCGCCAGCTCGGAGGCGCATCCGAGGCGGAGG 420  
QY CAGAGCTCGCGGCACTCTGCTCAGCCATGCGGAGCGCGCCGTCGACGTTAGACAGTCAA 480  
Db CAGAGCTCGCGGCACTCTGCTCAGCCATGCGGAGCGCGCCGTCGACGTTAGACAGTCAA 480  
QY ATCAGAGCTTCATGCTTAAGAGAGGAGTGAATGCAATTTGTCGCTTATTTGGAAGAAA 540  
Db ATCAGAGCTTCATGCTTAAGAGAGGAGTGAATGCAATTTGTCGCTTATTTGGAAGAAA 540  
QY TGAAGAGAGAGAGTTCGACGATGCAACCGAGTTCGTCATGTCATAGCAGATTAAGA 600  
Db TGAAGAGAGAGAGTTCGACGATGCAACCGAGTTCGTCATGTCATAGCAGATTAAGA 600  
QY AAATTTCTGCTGAATAAGGCGAGCGCGATTTGTTCCCTTTAAAGTTCGCGTTGATCAGT 660  
Db AAATTTCTGCTGAATAAGGCGAGCGCGATTTGTTCCCTTTAAAGTTCGCGTTGATCAGT 660  
QY CTGACCTGTCAATTAAGAAAGCTTGATGAGTTGACGAAGGACCTGGAATCCCTTCAGAAGG 720  
Db CTGACCTGTCAATTAAGAAAGCTTGATGAGTTGACGAAGGACCTGGAATCCCTTCAGAAGG 720  
QY AGAAGAGCGATCGGCTTAAGCAAGTGTAGTAACATTTGCAATTTCTTGCAATTCCTTATGTG 780  
Db AGAAGAGCGATCGGCTTAAGCAAGTGTAGTAACATTTGCAATTTCTTGCAATTCCTTATGTG 780  
QY AGGTGCTTGGCATAGATTTCAAGCAACAGATATATGAGGTGCACCTAGCTTGGACGAAG 840  
Db AGGTGCTTGGCATAGATTTCAAGCAACAGATATATGAGGTGCACCTAGCTTGGACGAAG 840  
QY CTGAAGGATCAAGAACTCTGAGCAACACTACAATTTGAGAGGCTTCTGCTGCGCGCAACA 900  
Db CTGAAGGATCAAGAACTCTGAGCAACACTACAATTTGAGAGGCTTCTGCTGCGCGCAACA 900  
QY GACTCGGTGAATGAGATCAAGAGTGCAGAACTTCAGATTTTCTGCTTACGATGC 960  
Db GACTCGGTGAATGAGATCAAGAGTGCAGAACTTCAGATTTTCTGCTTACGATGC 960  
QY TCGAGCTATGGAATCTCTGATGATCTCACTTGAAGAGCAGAGATGTTTCAGAAATATAA 1020  
Db TCGAGCTATGGAATCTCTGATGATCTCACTTGAAGAGCAGAGATGTTTCAGAAATATAA 1020  
QY CATGCAATTTGCTGCTTCAGAAACAGAGATTAACCTGAACCAACCTCTCCACAGATT 1080  
Db CATGCAATTTGCTGCTTCAGAAACAGAGATTAACCTGAACCAACCTCTCCACAGATT 1080  
QY TCCTGAAATTTGCTGAAATCTGAGGTTTAAAGCTTGAACCTGAACCAACCTCTCCACAGATT 1140  
Db TCCTGAAATTTGCTGAAATCTGAGGTTTAAAGCTTGAACCTGAACCAACCTCTCCACAGATT 1140  
QY AAGATCTTGTGTTTAAAAAGAGAGAGATTAAGAGAGATTAAGAGAGTGTCTCATCTTG 1200  
Db AAGATCTTGTGTTTAAAAAGAGAGAGATTAAGAGAGATTAAGAGAGTGTCTCATCTTG 1200

Db AAGATCTTGTGTTTAAAAAGAGAGAGATTAAGAGAGATTAAGAGAGTGTCTCATCTTG 1200  
QY TTGCGGAGGAGGTTATGCGAGAGGATTTAGCATTTGAGGCTATTGAAGCTGAGCATTTG 1260  
Db TTGCGGAGGAGGTTATGCGAGAGGATTTAGCATTTGAGGCTATTGAAGCTGAGCATTTG 1260  
QY ATCCCTCACTAGTACTTTGAAACAAATTTGAAGCTCACTTTGCAACAGTGAAGAGGAAGCTT 1320  
Db ATCCCTCACTAGTACTTTGAAACAAATTTGAAGCTCACTTTGCAACAGTGAAGAGGAAGCTT 1320  
QY TTAGCCGGAAGGATTTCTTGAGAAAGTTGAAAGATGCGCAAAATGCTTGTGAAAGGAAG 1380  
Db TTAGCCGGAAGGATTTCTTGAGAAAGTTGAAAGATGCGCAAAATGCTTGTGAAAGGAAG 1380  
QY CCTGGCTGGAAGATTAACAACAAAGATGATATCGTTACAAATGCTGGGAGGAGCAGCATC 1440  
Db CCTGGCTGGAAGATTAACAACAAAGATGATATCGTTACAAATGCTGGGAGGAGCAGCATC 1440  
QY TAAACACTAAAGAGGCTGAAAAGGCTCGTACTTTGGTCAACAAGATTCCTGGAATGGTAG 1500  
Db TAAACACTAAAGAGGCTGAAAAGGCTCGTACTTTGGTCAACAAGATTCCTGGAATGGTAG 1500  
QY ATGTTTGGAGAACAAAATTTGCTGATGGAAGGATGAAACGAGGAAAGAGGATTTCAAT 1560  
Db ATGTTTGGAGAACAAAATTTGCTGATGGAAGGATGAAACGAGGAAAGAGGATTTCAAT 1560  
QY ATGATGCTGTAGCTTTGCTCAATGCTTGAATATATGTTGCTCGTCAGGAGGAAG 1620  
Db ATGATGCTGTAGCTTTGCTCAATGCTTGAATATATGTTGCTCGTCAGGAGGAAG 1620  
QY AGCAAGAGAGAGAGAGAGAGGATCAGAAAGAGCTCCAGGATCAGCTCAAAAGCGGAGC 1680  
Db AGCAAGAGAGAGAGAGAGGATCAGAAAGAGCTCCAGGATCAGCTCAAAAGCGGAGC 1680  
QY AGSAGCTTTGTACGGATCAAAACCCAGTCCATCCAAAGCCCTAAGTACAAGAGGCAAC 1740  
Db AGSAGCTTTGTACGGATCAAAACCCAGTCCATCCAAAGCCCTAAGTACAAGAGGCAAC 1740  
QY CTAGGCACTCTATGGGTGGTGCBAACCGAAGGCTATCTTTGGTGGAGCCACCATGCAAC 1800  
Db CTAGGCACTCTATGGGTGGTGCBAACCGAAGGCTATCTTTGGTGGAGCCACCATGCAAC 1800  
QY CCCCAGAGACTGATATATCTGCAATTTCAAAGTCTGTTGCTGCCAAGAAAACTGAAAGAAA 1860  
Db CCCCAGAGACTGATATATCTGCAATTTCAAAGTCTGTTGCTGCCAAGAAAACTGAAAGAAA 1860  
QY TCGGCACTTTGTCCTTCTAGTAGTAGAGGTTTGGACATTTGCCGATTCCTATCAAGA 1920  
Db TCGGCACTTTGTCCTTCTAGTAGTAGAGGTTTGGACATTTGCCGATTCCTATCAAGA 1920  
QY AGTTGCTTTTCAATGCCAGTACTCTAGTGAGAGCGGAGACACCTCGTAAACCTTTTGCTC 1980  
Db AGTTGCTTTTCAATGCCAGTACTCTAGTGAGAGCGGAGACACCTCGTAAACCTTTTGCTC 1980  
QY AGATCAACACGAGGAAACAGTGTCTCGTCAAGCTGTCGCGCTTATCACCAATTAACATG 2040  
Db AGATCAACACGAGGAAACAGTGTCTCGTCAAGCTGTCGCGCTTATCACCAATTAACATG 2040  
QY AGGATGATGAGAACAGGACTCCGAAAGCATTTACAGCACTGAATCCCAAGACTCCGATGA 2100  
Db AGGATGATGAGAACAGGACTCCGAAAGCATTTACAGCACTGAATCCCAAGACTCCGATGA 2100  
QY CTGTTACGGCTCCCAATGCGAGTGGCAATGACTCCCTCTCGGCCAACAGGTTTTCAGCAA 2160  
Db CTGTTACGGCTCCCAATGCGAGTGGCAATGACTCCCTCTCGGCCAACAGGTTTTCAGCAA 2160  
QY CTCAGTTCCTTCTGTTTACGCAAGCCAGAGGTAAACATTTGAGGAGGACATCGACTACT 2220  
Db CTCAGTTCCTTCTGTTTACGCAAGCCAGAGGTAAACATTTGAGGAGGACATCGACTACT 2220  
QY CCTTTGAAGAAAGCGGCTCGCCATCTATCTGGCCAGGCAAAATGTTTAACTGTTGATCA 2280  
Db CCTTTGAAGAAAGCGGCTCGCCATCTATCTGGCCAGGCAAAATGTTTAACTGTTGATCA 2280

[illegible]





Db 983 TGCAGAGCTTCAAGATTTGGCTCTAGCATGCTTGAATCTCATGGATACAC 1042  
Qy 988 CACTTGAAGAGCAGCATGTTTCAAGATTAACATGCAATATGCTGCTTCAAGACAG 1047  
Db 1043 CACTTGAAGAGCAGCATGTTTCAAGATTAACATGCAATATGCTGCTTCAAGACAG 1102  
Qy 1048 AGATACTGAACCAACACCTCTCCACAGATTTCTGCAATATGTCGAATCTGAGGTGT 1107  
Db 1103 AATAACTGAGCTTAACAGCTCTCTACTGCTTCTAGCTACGTGGAATCTGAAGTTT 1162  
Qy 1108 TAAGGCTTGAACAACTTGAAGCAAGTAAAGATGAAAGATCTTGTTTTAAAAAGAAAGCAG 1167  
Db 1163 TAAGACTTGAACAGCTTAAAGCGAGCAAGATGAAGACCTTGTCTTAAAAAGAAAGACAG 1222  
Qy 1168 AACTAGAAGCATGAAGACGTGCTCATCTGTTGGCGAGGAAGGTATGCAAGGACT 1227  
Db 1223 AACTGGAAGATCATAGGACGCTGCTCATTTGATTGGCGAGGAAGGCTATGCAAGTGAAT 1282  
Qy 1228 TTAGCATTTGAAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTG 1287  
Db 1283 TTAGCGATGAGGCTATTGAGCGAGGAGCTGTTGATCTTCTGCTGCTGGAACAAATTTG 1342  
Qy 1288 AAGCTCACATTTGCAACAGTGAAGAGGAAGCTTTTAGCCGGAAGGATATTTCTTGAGAAAG 1347  
Db 1343 AGGCTCACATTTGCCACAGTGAAGAGGAAGGCTTTTAGCAGGAAGGATATTTCTCGAAGAG 1402  
Qy 1348 TTGAAGATGGCAAAATGTTTGAAGAGAGGAGGCTGGCTGGAGAGTATTAACAAAGATG 1407  
Db 1403 TTGAAGATGGTGAATGCAATGTTGAGGAGGAAGCTTTGGCTGGAAGATTAACAAAGATG 1462  
Qy 1408 ATAATCGTTACATGCTGGGAGGGAGGACATCTAACACTAAAGAGGCTGAAAGGCTC 1467  
Db 1463 ACAATCGTTATATGCGGAGGGAGGGGCCACTTGACACTCAAAAGGCTGGAAGGCTC 1522  
Qy 1468 GTACTTTGGTCAACAAAGATTTCTGGAATGGTAGATGTTTGAAGCAAAATTTGCTGCAAT 1527  
Db 1523 GTATTTTGGTTAACAAGATCCAGGACTGGTAGATGTTTAAACCAACAAATTTGAGCCT 1582  
Qy 1528 GGAATAATGAACAGAGGAAGGAGGATTTACATATGATGGTGTAGCCTTCTGTAATGC 1587  
Db 1583 GGGAGGCGAGAAAGAGGAA---AGGAATTTACATACGACGCTGTCGCGCTTATGTAATGC 1639  
Qy 1588 TTGATGATATATGTTGCTGTCAGGAGAAAGCAAGCAAGAGAGAGCAAGGATC 1647  
Db 1640 TTGAAGAGTACATGCTGTTGCGCAGGAGGAAGGCTACAGAAGAGGCAAGGATC 1699  
Qy 1648 AGAAGAGCTCCAGGATCAGCTCAAAAGCGGAGGAGGAGCTTTGTACGGATCAAAACCCA 1707  
Db 1700 AGAAGAAATCCAGGATCAACTCAAGCTGAGCAGGAAGGCTCTATGGATCAAAACCAA 1759  
Qy 1708 GTCATCCAAAGCCCTTAAGTAAAGAGGCACTTAGGCACTCTATGGGTGGTGAACACC 1767  
Db 1760 GTCCTTTAAGCCTCAAAAGCAAGCAAGAGGCGCTAGGCACTCCATGGGTGGTGGCAACC 1819  
Qy 1768 GAAGGCTATCTTGGTGGAGCCACCATGCAACCCCGAGAGCTGATATACTCATTTCAA 1827  
Db 1820 GAAGACTGTCTTGGTGGAGCCACTATGCAAGCCCCCAAAACAGATATACTGCACTCCA 1879  
Qy 1828 AGTCTGTTGCTGCTGCCAAGAAACTGAAGAAATCGGCACTTTGTCCCTTAGTAGTAGTA 1887  
Db 1880 AGACAGCTCGTGGCGCCCAAGAGGCTGAAGATTTGGGCGCTTTATCTCC---TAGTAGTA 1936  
Qy 1888 GAGGTTTGGACATTTGCCGATTTGCTATCAAGAGTTGCTTTTCAATGCCAGTACTCTAC 1947  
Db 1937 GAGGCTTGGACATTTGCCGCTTTCTCCCATCAAGAAGTTATCTTTCAACGCAAGCACTTTGC 1996  
Qy 1948 GTGAGCGGAGACACTCGTAAACCTTTTGTCTGATGCACACCGAGAAACAGTGTCTCGT 2007  
Db 1997 GAGAGGAGAGAAACACCGCGAAGGCTTTTGGCCAGATCATGCCAGGAAACAGCTCTCAC 2056  
Qy 2008 CGAGCGCTGTGGCGCTTATACCAATAACACTGAGGATGATGAGAAACAGGACTCCGGAAGA 2067

Db 2057 CGATGCTTACGCGGCCCATCTCCAGGCCACAGAGGAAGAGAACAAAACCCCCAGACAT 2116  
Qy 2068 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTTACGCGTCCAATGCAATGGCAA 2127  
Db 2117 TTGTAGGAGGCTCAATCGAAAGCGCGGACAGTACGCTCTCTATGACAGATGGCGA 2176  
Qy 2128 TGACTCCCTCTCTGCGCAACAGGTTTCAGCAACTCCAGTTCCTTCTTGTACGACACAGC 2187  
Db 2177 TGACACCAAGCGGTGGCTTAACAAGGCTATAGCCACTCCTGCAACCCCTTTTCCAGGAGAGG 2236  
Qy 2188 CAGA---GGTAACATTGAGGAGGACATCGACTACTCTCTTTGAAGAAAGGCGGCTCGCCA 2244  
Db 2237 CAGATTCGCCAGACTGCGAGCGGACATCGACTACTCGTTGGAAGAGGCGGCTCGCTG 2296  
Qy 2245 TCTATCTGCGCCAGGCAAAATGGTTTAACTGTTGATCAATTTATGTAAGTAAATCTG 2304  
Db 2297 TTTACTGCGCCAGCAAGTGGCTTAA-ACTCGTCACTTACGTAAGTGAAGCTGAACCTG 2355  
Qy 2305 ACTGCATTTTCTTGTGCGTGGCCATTTG 2331  
Db 2356 ACTGCAATTTCTCGTCAGAGGCCAATG 2382

## RESULT 4

US-10-619-685-3  
; Sequence 3, Application US/10619685  
; Publication No. US20040096875A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Resources, Ministry of Agriculture,  
; APPLICANT: Forestry and Fisheries  
; APPLICANT: Japan Science and Technology Corporation  
; TITLE OF INVENTION: A NOVEL GENE FOR CONTROLLING LEAF SHAPES  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/619,685  
; CURRENT FILING DATE: 2003-07-15  
; PRIOR APPLICATION NUMBER: US/09/667,475D  
; PRIOR FILING DATE: 2002-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 4574  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-10-619-685-3

Query Match 41.4%; Score 1021.2; DB 16; Length 4574;  
Best Local Similarity 69.6%; Pred. No. 9.1e-278;  
Matches 1974; Conservative 0; Mismatches 3; Indels 861; Gaps 9;

Qy 469 TTAGACAGTCAAAATCAGAAAGCTTTCATGGCTTAAGAGAGAGTGAATGCAATTTGTTCCGT 528  
Db 1177 TTTTGCAGTCAAAATCAGAAAGCTTTCATGGCTTAAGAGAGAGTGAATGCAATTTGTTCCGT 1236  
Qy 529 ATTTGGAAGAAATGAAGAAAGGCTCGAAGTGAAGCAAGTGAATTTGTTTCAATGTCATAG 588  
Db 1237 ATTTGGAAGAAATGAAGAAAGGCTCGAAGTGAAGCAAGTGAATTTGTTTCAATGTCATAG 1296  
Qy 589 ACAGAGTAAAGAAATTTTCGTCGTAATAAGGCGAGCGAGTGAATTTGTTTCCCTTTAAAGTTC 648  
Db 1297 AGCAGATTAAGAAATTTTCGTCGTAATAAGGCGAGCGAGTGAATTTGTTTCCCTTTAAAGTTC 1356  
Qy 649 CGGTTGATCAGTCTGACCTGTCTATTAAGAAAGCTTGAATGAGTGAAGCAAGGACTCGAAT 708  
Db 1357 CGGTTGATCAGTCTGACCTGTCTATTAAGAAAGCTTGAATGAGTGAAGCAAGGACTCGAAT 1416  
Qy 709 CCCTTCAGAGGAGA----- 723  
Db 1417 CCCTTCAGAGGAGGAGGCTCATCATCAATAACATCTTTATCCATTTTCCAGGACTCAT 1476  
Qy 724 ----- 723  
Db 1477 GTTGTGATCGTGTCTCTATCTATCAAGAAATCTTTTTCATTTCTTGTATATAAATCTCACTA 1536

QY 724 -----AGCGGATCGCTAAAGCAAGTGATAGAAATTG 758  
DB 1537 TGCCATATACATGTTTGTCTCACAGACGATCGCTTAAGCAAGTGATAGAAATTG 1596  
QY 759 AATTCTTTCATTCCTTATGTGAGGTGCTTGCATAGATTTCAAGCAACAGTATATGAG 818  
DB 1597 AATTCTTTCATTCCTTATGTGAGGTGCTTGCATAGATTTCAAGCAACAGTATATGAG 1656  
QY 819 GTGACCCCTAGCTTGGACGAAGCTGAAGATCAAGAAACCTGAGCAACATCAATTTGAG 878  
DB 1657 GTGACCCCTAGCTTGGACGAAGCTGAAGATCAAGAAACCTGAGCAACATCAATTTGAG 1716  
QY 879 AGGCTTGCTGCTGCCGCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAA----- 933  
DB 1717 AGGCTTGCTGCTGCCGCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGGTC 1776  
QY 934 ----- 933  
DB 1777 AGCATTTGCCCTGTACCATTGTAGAGGTATCAATGAACACTTTTCAGTCTTTAACTTGGTTAA 1836  
QY 934 -----AGCTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGG 981  
DB 1837 TCTGATTTGCGCAGCTTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGG 1896  
QY 982 ATACTCCACTTGAAGAGCAGCAGATGTTTCAGAAATATAACATGCAATATTGCTGCTTCAG 1041  
DB 1897 ATACTCCACTTGAAGAGCAGCAGATGTTTCAGAAATATAACATGCAATATTGCTGCTTCAG 1956  
QY 1042 AACAGAGATTAATCTGAACCAACACACCTCTCCACAGATTTTCCTGAATAT----- 1091  
DB 1957 AACAGAGATTAATCTGAACCAACACACCTCTCCACAGATTTTCCTGAATATGTAATTTATC 2016  
QY 1092 ----- 1091  
DB 2017 ATCACTGAGATTGCAAAAATTATATGTTGCTACTGTTTATATTTTCAATAAGATATGAAT 2076  
QY 1092 -----GTGGAATCTGAGGTGTTAAGGCTTGAAACAACTG 1124  
DB 2077 GTTCATCGACTACTATTATACCTGTAGTCTGAACTCTGAGGTGTTAAGGCTTGAAACAACTG 2136  
QY 1125 AAGCAAGTAAGATGAAGATCTTGTGTTTAAAAAGAAAGCAAGACTAGAGAGCATAGA 1184  
DB 2137 AAGCAAGTAAGATGAAGATCTTGTGTTTAAAAAGAAAGCAAGACTAGAGAGCATAGA 2196  
QY 1185 AGACGTGCTCATCTTGTGTCGAGAGAGGTTATGCAAGAGGAGTTAGCATTGAAGCTATT 1244  
DB 2197 AGACGTGCTCATCTTGTGTCGAGAGAGGTTATGCAAGAGGAGTTAGCATTGAAGCTATT 2256  
QY 1245 GAAGCT----- 1250  
DB 2257 GAAGCTGGTAAGATACTCTCTGCTTACTGCTTTTATTGTGCTGACAAAGTCATACCA 2316  
QY 1251 -----GGAGCTATGATCCCTCACTAGT 1273  
DB 2317 GACAGAGTTCAATACCTGCTGCTGTTCTGTTTCGAGGAGCTATTGATCCCTCACTAGT 2376  
QY 1274 ACTTGAAACAAATTGAAGCTCAATTGCAACAGTGAAGAGGAGCTTTTAGCCGAAGGA 1333  
DB 2377 ACTTGAAACAAATTGAAGCTCAATTGCAACAGTGAAGAGGAGCTTTTAGCCGAAGGA 2436  
QY 1334 TATCTTGAGAAAGTTGAAGATGCAAAATGCTTTGTGAAGAGGAGGCTGCTGGAAGA 1393  
DB 2437 TATCTTGAGAAAGTTGAAGATGCAAAATGCTTTGTGAAGAGGAGGCTGCTGGAAGA 2496  
QY 1394 TTACAACAAA----- 1403  
DB 2497 TTACAACAAAGTATGAGTGTAGCTGAAGCTACGTTGTTTGTATATTTTGTAGCAAA 2556  
QY 1404 -----GATGATAATCGTTACAATGC 1423  
DB 2557 TAATGTGTAAGTATATCTCTGGCTTTGGCTTTTGTAGGATGATAATCGTTACAATGC 2616  
QY 1424 TGGAGGGGAGCACATCTTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTGGTCAACAA 1483

DB 2617 TGGAGGGGAGCACATCTTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTGGTCAACAA 2676  
QY 1484 GATTCCT----- 1490  
DB 2677 GATTCCTGGTAATGTTACTCAATGATTTATGTTTGGAACTTCTCTTATCAAGTGCATAT 2736  
QY 1491 ----- 1490  
DB 2737 TTAATTTACAATTTTAACTCTTGCCATTACTAACAATCTGATATCTCTGCTGATTTGTGCTG 2796  
QY 1491 -----GGAATGCTAGATGTTTGGAAACAAAATTTGCTCATGGAATAATGAACAGGAAA 1546  
DB 2797 AGCAGGAATGCTAGATGTTTGGAAACAAAATTTGCTCATGGAATAATGAACAGGAAA 2856  
QY 1547 GGAGGAATTCACATATGATGCT----- 1569  
DB 2857 GGAGGAATTCACATATGATGCTGATGTTTCTTACTCTTACACATTTACATTTGATCGGGT 2916  
QY 1569 -----GTTAGCCTTTCTGTC 1583  
DB 2917 CTATTTTGTGTTTCTGCTGAAGTGCCTTCTTGGCAATTTCTTACAGGTTAGGCTTTCTGTC 2976  
QY 1584 ATGCTTGTATGAATATATGTTTCTGCTGAGAGAAAGCAAGAGAAAGCAAGCAAAAG- 1642  
DB 2977 ATGCTTGTATGAATATATGTTTCTGCTGAGAGAAAGCAAGAGAAAGCAAGCAAAAG 3036  
QY 1643 ----- 1642  
DB 3037 GTATTATGCTCTCGCCTAATATTCAATGATTTGTTCTAAATCATCTTTTACCTTCTGTGAA 3096  
QY 1643 -----GGATCAGAAGAGCTCCAGATCAGATCAGCTCAAG 1674  
DB 3097 TACCTCTAATACTTGAATATATACCTCGAGGATCAGAAGAGCTCCAGATCAGCTCAAG 3156  
QY 1675 CGGAGCAGGAAGCTTTGTAGCGATCAAAACCCAGTCCATCCAGCCCTCAAGTCAAAAGA 1734  
DB 3157 CGGAGCAGGAAGCTTTGTACGGATCAAAACCCAGTCCATCCAGCCCTCAAGTCAAAAGA 3216  
QY 1735 AGGCACCTAGGACCTCTATGCGGTGTCAAACCCAGGCTATCTCTTGGTGGAGCCACCA 1794  
DB 3217 AGGCACCTAGGACCTCTATGCGGTGTCAAACCCAGGCTATCTCTTGGTGGAGCCACCA 3276  
QY 1795 TGCAACCCCGAAGACTGATATCTGCAATCAAAAGTCTGTTGCTGCTGCCAAGAAAACCTG 1854  
DB 3277 TGCAACCCCGAAGACTGATATCTGCAATCAAAAGTCTGTTGCTGCTGCCAAGAAAACCTG 3336  
QY 1855 AAGAAATCGGCACCTTTGTCCTCTAGTA----- 1881  
DB 3337 AAGAAATCGGCACCTTTGTCCTCTAGTAAGCCCTACTAGTATCATGTGTCGATATATTTTC 3396  
QY 1882 -----GTAGTA 1887  
DB 3397 TTTTTCCTCTTATTTTCACTTGAAACATATGTTTAACTCAAGCAAAACATATCAAGTAGTA 3456  
QY 1888 GAGGTTTGGACATTTGCCGATTTGCTATCAAGAGTTGTTCTTTCAATGCCAGTACTCTAC 1947  
DB 3457 GAGGTTTGGACATTTGCCGATTTGCTATCAAGAGTTGTTCTTTCAATGCCAGTACTCTAC 3516  
QY 1948 GTGAGACGGAGACACCTCGTAAACCTTTTGTCTAGATCAACCAAGGAAAACAGTGTCTCGT 2007  
DB 3517 GTGAGACGGAGACACCTCGTAAACCTTTTGTCTAGATCAACCAAGGAAAACAGTGTCTCGT 3576  
QY 2008 CGAGCCTGTCGCGCTTATCAACCAATCAACTGAGGATGATGAGAACAGGACTCCGAGA 2067  
DB 3577 CGAGCCTGTCGCGCTTATCAACCAATCAACTGAGGATGATGAGAACAGGACTCCGAGA 3636  
QY 2068 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTTACGGCTCCCAATGAGATGCAAA 2127  
DB 3637 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTTACGGCTCCCAATGAGATGCAAA 3696  
QY 2128 TGACTCCTCTCTGCGCAACAAAGTTTCAGCAACTCAGTTTCCCTTGTGTTTACCAAGC 2187

Db	3697	TGACTCCTCTCTGGCCAAACAAGGTTTCAGACACTCCAGTTTCCCTTGTTTACGACAAGC	3750
Qy	2188	CAGAGGTAAACATTGACAGGAGGACATCGACTACTCTCTTTGAAGAAAGCGCGCTCGCCATCT	2247
Db	3757	CAGAGGTAAACATTGACAGGAGGACATCGACTACTCTCTTTGAAGAAAGCGCGCTCGCCATCT	3816
Qy	2248	ATCTGGCGCAGGCAAAATGGTTTAACTGTTTGATCAATTTATGTACGTAGTTCGAAATCTGACT	2307
Db	3817	ATCTGGCGCAGGCAAAATGGTTTAACTGTTTGATCAATTTATGTACGTAGTTCGAAATCTGACT	3876
Qy	2308	GCAATTTCTTGTGGTGGCCAAATGCGGTATGTTGGTCAACAATAGTCGGCCCTTTCCAGTAG	2367
Db	3877	GCAATTTCTTGTGGTGGCCAAATGCGGTATGTTGGTCAACAATAGTCGGCCCTTTCCAGTAG	3936
Qy	2368	CACATTTCTCATTTTACGTGCAATTTGTTTTAAATGTTTCTACAACACAGTAAACAGCTCTCAT	2427
Db	3937	CACATTTCTCATTTTACGTGCAATTTGTTTTAAATGTTTCTACAACACAGTAAACAGCTCTCAT	3996
Qy	2428	ACATTAGCTTGCTCACTA	2445
Db	3997	ACATTAGCTTGCTCACTA	4014

```

RESULT 5
US-10-437-963-95930/c
; Sequence 95930, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95930
; LENGTH: 784
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94075C.1
US-10-437-963-95930

```

Query Match	27.3%	Score 674.4	DB 17	Length 784
Best Local Similarity	97.7%	Pred. No. 6e-180		
Matches 717	Conservative 0	Mismatches 11	Indels 6	Gaps 3
Qy	1714	CCAAGCCCCTAAGTACAAAGAAGGCACCTAGGCACCTCTATGGGTGGTGCACCAACCGAAGGC	1773	
Db	778	CCAAGCCCCTAAGTACACGGAAGGCACC-AGGCACCTCTATGGGTGGTGCACCAACCGAAGGC	720	
Qy	1774	TATCTCTTGGTGGAGCCACCATGCAACCCCGAGACGTATATCTGCAATTCAAAGTCTG	1833	
Db	719	TATCTCTTGGTGGAGCCACCATGCAACCCCGAGACGTATATCTGCAATTCAAAGTCTG	660	
Qy	1834	TTCGTGCTGCCAAGAAAACTGAAAGAAATCGGCACCTTTGTCGCCCTAGTAGTAGAGAGTT	1893	
Db	659	TTCGTGCTGCCAAGAAAACTGAGAAATCGGCACCTTTGTCGCC- - -TAGTAGTAGAGTT	603	
Qy	1894	TGGACATTTGCCGATTGCCATCAAGAAGTTGTCTTTCAATGCCAGTACTCTACGTGAGA	1953	
Db	602	TGGACATTTGCCGATTGCCATCAAGAAGTTGTCTTTCAATGCCAGTACTCTACGTGAGA	543	
Qy	1954	CGGAGACCTCGTGTAAACCTTTTGTCTAGATCACACCGAGGAAAAGTGTCTCTGTCGACGC	2013	
Db	542	CGGAGACCTCGTGTAAACCTTTTGTCTAGATCACACCGAGGAAACAGTGTCTCTGTCGACGC	483	

2014	Qy	CTGTGGCCCTTATCCCAATTAACACTGAGATGATGAGAACAGGACTCCGGAACACATTTA	2071
482	Db	CTGTGGCCCTTATCCCAATTAACACTGAGATGATGAGAACAGGACTCCGGAACACATTTA	423
2074	Qy	CAGCACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCGAGATGGCAATGACTC	2133
422	Db	CAGCACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCGAGATGGCAATGACTC	363
2134	Qy	CCTCTCTGGCCAAACAAGGTTTCAGCAACTCCAGTTTCCCTTGTATTACGACAAGCCAGAGG	2193
362	Db	CCTCTCTGGCCAAACAAGGTTTCAGCAACTCCAGTTTCCCTTGTATTACGACAAGCCAGAGG	303
2194	Qy	TAAcATTGcAGGAGGAcATcGcAcTAcTcCTTTTcAAGAAAGcGGcCTcGcCAcTcTAcTcTcGg	2253
302	Db	TAAcATTGcAGGAGGAcATcGcAcTAcTcCTTTTcAAGAAAGcGGcCTcGcCAcTcTAcTcTcGg	243
2254	Qy	CCAGGCAAAATGGTTTTAACTGTGTTGATCAAAATTTATGTACGTAGTTGAAATCTGACTGCATTT	2313
242	Db	CCAGGCAAAATGGTTTTAACTGTGTTGATCAAAATTTATGTACGTAGTTGAAATCTGACTGCATTT	183
2314	Qy	TCTTGTcGGTGGcCATGTCGTATGTTGGTCAAAcATAGTCGGcCTTTCCAGTAgcAcTAT	2373
182	Db	TTTGTcGGTGGcCATGTCGTATGTTGGTCAAAcATAGTCGGcCTTTCCAGTAgcAcTAT	123
2374	Qy	TCTGATTTACTGCAAAATGTTTTAAATGTTTTT - -CTACAACcAGTAAAAcAGcCTcTATAcAT	2431
122	Db	TTTGAATTTACTGCAAATGTTTTAAATGTTTTTCTACAACcAGTAAAAcAGcCTcTATAcAT	63
2432	Qy	TAGCTTGCTCACTA 2445	
62	Db	TAGCTTGCTCACTA 49	

```

RESULT 6
US-10-437-963-81367
; Sequence 81367, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with Rice
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81367
; LENGTH: 709
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(709)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80900C.1
US-10-437-963-81367

```

[illegible]

QY 64 CTCCTCCGCTCTCTCGCATCTAGGCTCCGATCGCGCGGACCCAGCAGAAATCCGCC 123  
Db 131 CTCCTCCGCTCTCTCGCATCTAGGCTCCGATCGCGCGGACCCAGCAGAAATCCGCC 190  
QY 124 GCCCGCTCTCGGCTCCCGCTCGAGGACCGCGCGGCGGAGGCGCTAGTGT 183  
Db 191 GCCCGCTCTCGGCTCCCGCTCGAGGACCGCGCGGCGGAGGCGCTAGTGT 250  
QY 184 CTTTCGACCTTCGCGATGAGTACGCGGTGAAGGACCGCTTCACAGATGTCGAGCAT 243  
Db 251 CTTTCGACCTTCGCGATGAGTACGCGGTGAAGGACCGCTTCACAGATGTCGAGCAT 310  
QY 244 GCGATTTCGCTTCTACTGGAGCTCAATGTGATTGGGATGAGTCCGTCGAGCCCGACGA 303  
Db 311 GCGATTTCGCTTCTACTGGAGCTCAATGTGATTGGGATGAGTCCGTCGAGCCCGACGA 370  
QY 304 CGAGGACAGGATGCTCTGGAGCTCGAGCAGAGTGCCTGAGTCTACAGGCGGAAGG 363  
Db 371 CGAGGACAGGATGCTCTGGAGCTCGAGCAGAGTGCCTGAGTCTACAGGCGGAAGG 430  
QY 364 TCGACAGCGGAACCGGAGCCCGCCAGCTCGGAAGGCCATCCCGAGGCGGAGGCG 423  
Db 431 TCGACAGCGGAACCGGAGCCCGCCAGCTCGGAAGGCCATCCCGAGGCGGAGGCG 490  
QY 424 AGCTCCCGGCTCTCTCTCAGCCATGGCGAGCCCGCGTGCACTTACAGAGTCAAAATC 483  
Db 491 AGCTCCCGGCTCTCTCTCAGCCATGGCGAGCCCGCGTGCACTTACAGAGTCAAAATC 550  
QY 484 AGAAGCTTCATGCTTAAGAGAGGAGTTGAATGCAATTTGTCCTTTAAA 543  
Db 551 AGAAGCTTCATGCTTAAGAGAGGAGTTGAATGCAATTTGTCCTTTAAA 609  
QY 544 AAAAGAAAAGTTCGAACGATGGAACCGTCTTGTTCATGTCATAGCAGAGATTAAAGAAA 603  
Db 610 -AAAGAAAAGTTCGAACGATGGAACCGTCTTGTTCATGTCATAGCAGAGATTAAAGAAA 668  
QY 604 TTTCTGTGAATAAGGCGGAGCCGATTTGTTCCCTTTAAA 644  
Db 669 TTTCTGTGAATAAGGCGGAGCCGATTTGTTCCCTTTAAA 709

## RESULT 7

US-10-437-963-98451  
; Sequence 98451, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 98451  
; LENGTH: 2045  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MKT4530\_96356C.1  
US-10-437-963-98451

Query Match 22.8%; Score 563; DB 17; Length 2045;  
Best Local Similarity 76.6%; Pred. No. 4.4e-148;  
Matches 689; Conservative 0; Mismatches 210; Indels 0; Gaps 0;  
QY 198 ATGAGTAGCGGTGAAGACGAGCTTCCAGCATGTCCGACATGCGATTTCGCTTCTA 257

Db 1 AGGAGCGCAATCCGAGGAGCGCTCTCTCCGATGAAGACCATATGCGATCCCTCATG 60  
QY 258 CTGGAGCTCAATGTGATTGGGATGAGGTCCGTTAGCCCGACACGACGAGGACAGGATG 317  
Db 61 CACGAGCTCCAGGTCTCTGGGACGAGGTCCGGGAGCCCGGCGGCGGACAGGATG 120  
QY 318 CTGCTGGAGCTCGAGCAGGAGTGCCTGGAGGTCTACAGGCGGAAGGTGCACACGCGCAAC 377  
Db 121 CTGCTGGAGCTCGAGCAGGAGTGCCTGGAGGTCTACAGGAGGAGGTTGGACACGCGCAAC 180  
QY 378 CGSAGCCGCGCCAGCTCGGGAAGGCCATCCCGAGGCGGAGGAGCTCCGCGGATC 437  
Db 181 CGCTCCGCGCCAGCTCGGCGAGCGATCGCTCAATATGAAGCTGAACCTCGCGCCATC 240  
QY 438 TGCTCAGCATGCGGCGGAGCCCGCTGCAGCTTAGACAGTCAATCAGAACTTCATGGC 497  
Db 241 TGCTCCGCGCATTCGCGAGACGACGGTCCAGCTAGGCGAGTCTAATCAGAAAGCATGTGGA 300  
QY 498 TTAAGAGAGGAGTTGAATGCAATTTGTTCCGTTATTGGAAGAAATGAAAAGAAAAGGTC 557  
Db 301 TTACGGGATGAGCTCGGTGCAATATTACCATACCTAGAGAGATGAGAGGAGGAGGTT 360  
QY 558 GAACGATGGAACCGATTGTTTCATGTCTAGAGCAGATTAGAAAAATTTCTCTGAATA 617  
Db 361 GAAAGATGGAACCGATTGTTTCATGTCTAGAGAGGATTAAGAAAGATCTCATCTGAGATA 420  
QY 618 AGCCAGCGGATTTGTTCCCTTTAAAGTTCCGGTTCAGTCTGAGCTGTCTATTAGA 677  
Db 421 AGCCAGCAAAATTTGACCCCTTTTAAAGTGTCTGGATCAATCTGATCTATCATTAAGA 480  
QY 678 AAGCTTGATGAGTTGACGAGGACCTGGAATCCCTTCAGAGGAGAGAGCGATCGGCTA 737  
Db 481 AAGCTTGAGAGTTAAGGTGGAGCTGAGTCCCTTGAAAGGAGAGGAGGAGGTA 540  
QY 738 AAGCAAGTATGAGAAATTTGAATTTCTTTCATTTCTTATGTAGGAGTCTGGCATAGAT 797  
Db 541 AAGCAAGTATGAGAAATTTGAAGACTTTAAGACTTTTATGTGTAGTACTTGTGTGAC 600  
QY 798 TTCAGCAACACAGTATATGAGTGCACCTGCTTGCAGCAAGCTGAGGATCAAGAAC 857  
Db 601 TTCAGCAACACATATCTGAAATACCCCTAGTCTTGTATGAGCTGAGGCGCAAGGAT 660  
QY 858 CTGAGCAACACTACAATTTGAGAGGCTTGTCTGCTCCGCAACACAGACTCGCTGAAATGAAG 917  
Db 661 ATAAGCAACACTACAATTTGAGATGCTAGCATGGGCGATTCAGAGACTTCGTGAAACAAA 720  
QY 918 ATCCAAAGGATGCAAAAGCTTCAAGATTTTGTCTTCTAGCATGCTCGAGCTATGGAATCTC 977  
Db 721 ATGCAGAGGATGCAAAAGCTTCAAGATCTTGCATCTACCTTGTCTAGAACTATGGAATCTA 780  
QY 978 ATGGATCTCCACTTTGAAGGCGAGCATGTTTTCAGAAATATAACATATTTGCTGCT 1037  
Db 781 ATGATAGCGCATTTTGAAGAGCAGCAGCATACAGAAATATAACATATTTGCTGCT 840  
QY 1038 TCAGAAACAAGAGATAAATCTGAACCAACACCCCTCTCCACAGATTTCTGGAATATGTCGA 1096  
Db 841 TCAGAGGCTGAATTAACAGAACAGAACACCCCTTTCCATTTGAGTTTCTTCACTATGATGA 899

## RESULT 8

US-10-437-963-32360  
; Sequence 32360, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping

;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
;; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

;; CURRENT APPLICATION NUMBER: US/10/437,963

;; CURRENT FILING DATE: 2003-05-14

;; NUMBER OF SEQ ID NOS: 204966

;; SEQ ID NO 32360

;; LENGTH: 530

;; TYPE: DNA

;; ORGANISM: Oryza sativa

;; FEATURE:

;; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_36577C.1

US-10-437-963-32360

Query Match 20.5%; Score 506.6; DB 17; Length 530;

Best Local Similarity 99.0%; Pred. No. 1.7e-132;

Matches 520; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1163 AGCAGAACTAGAGACATAGAGACGCTGCTCATCTTGTGGCGAGGAAGGTTATGCAGA 1222

DB 6 AGGAGAACTTGAAGACATAGAGACGCTCATCTTGTGGCGAGGAAGGTTATGCAGA 65

QY 1223 GGAGTTTACATTTGAAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACA 1282

DB 66 GGAGTTTACATTTGAAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACA 125

QY 1283 AATTGAAGCTCACAATGCAACAGTGAAGAGGAGCTTTAGCCGGAAGGATATCTTGA 1342

DB 126 AATTGAAGCTCACAATGCAACAGTGAAGAGGAGCTTTAGCCGGAAGGATATCTTGA 185

QY 1343 GAAAGTTGAAGAGTGCAGAAATGCTT-GTGAAGAGGAAGCTGGCTGGAAGATTACAACA 1401

DB 186 GAAAGTTGAAGAGTGCAGAAATGCTTGGGAGAGGAGGCTGGCTGGAAGATTACAACA 245

QY 1402 AAGATGATAATCGTTACAATGCTGGAGGAGGAGCAATCTAACACTAAAGAGGGCTGAAA 1461

DB 246 AAGATGATAATCGTTACAATGCTGGAGGAGGAGCAATCTAACACTAAAGAGGGCTGAAA 305

QY 1462 AGGCTGCTACTTGGTCAACAAGATTCCTGGATGCTAGATGTTTGAAGAACAAAATTG 1521

DB 306 AGGCTGCTACTTGGTCAACAAGATTCCTGGATGCTAGATGTTTGAAGAACAAAATTG 365

QY 1522 CTCGATGGAAAAATGAACGAGGAAGAGGATTTACATATGATGGTGTAGCCTTTGCT 1581

DB 366 CTCGATGGAAAAATGAACGAGGAAGAGGATTTACATATGATGGTGTAGCCTTTGCT 425

QY 1582 CAATGCTTGATGATATATGTTCTGTCGAGAGAAAGCAAGAGAAAGAGACAAA 1641

DB 426 CAATGCTTGATGATATATGTTCTGTCGAGAGAAAGCAAGAGAAAGAGACAAA 485

QY 1642 GGGATCAGAGAAGCTCCAGGATCAGCTCAAGCGGAGCAGGAAG 1686

DB 486 GGGATCAGAGAAGCTCCAGGATCAGCTCAAGCGGAGCAGGAAG 530

## RESULT 9

US-10-424-599-126041

;; Sequence 126041, Application US/10424599

;; Publication No. US20040031072A1

;; GENERAL INFORMATION:

;; APPLICANT: La Rosa Thomas J

;; APPLICANT: Kovalic David K

;; APPLICANT: Zhou Yihua

;; APPLICANT: Cao Yongwei

;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

;; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

;; CURRENT APPLICATION NUMBER: US/10/424,599

;; CURRENT FILING DATE: 2003-04-28

;; NUMBER OF SEQ ID NOS: 285684

;; SEQ ID NO 126041

;; LENGTH: 2150

;; TYPE: DNA

;; ORGANISM: Glycine max

;; FEATURE:

;; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_84823C.1

US-10-424-599-126041

Query Match 19.0%; Score 467.8; DB 16; Length 2150;

Best Local Similarity 58.8%; Pred. No. 4.5e-121;

Matches 887; Conservative 0; Mismatches 607; Indels 15; Gaps 4;

QY 240 ACATCGGATTGCTTCTACTGAGCTCAATGTGATTTTGGGATGAGGTCGGTGAGCCGAC 299

DB 139 ACATGTGTTCTTCTGTTTAAAAAGCTTCAGGAAATATGGATGAGGTTGGTGAGCGAC 198

QY 300 ACGCAGGAGCAGGATGCTGCTGAGCTCGAGCAGGAGTGCTCGAGGCTCTACAGCGG 359

DB 199 GAGCAACGAGACAAGATGCTTCTCAGTTAGACGAGGAGTGCTTGGATGTGTACAAGAGA 258

QY 360 AAGTTCGACAGCGCAACCGGAGCGCCCGACGCTCGGGAAGGCCATCGCCGAGCGCAG 419

DB 259 AAGTTCGACAGCGCTGCAAAAGTCAAGGGCGCAGCTACTTCAAGCTCTGTCTGATTAAG 318

QY 420 GCAGAGCTCGCCGGCATCTGCTCAGCCATGGCGCAGCGCCCGTGCACGTTAGACAGTCA 479

DB 319 CTTGAGCTTTCACCTCTTCTATCAGCACTTGGAGAAAGAGCTT---TGCTGGAATTCCT 375

QY 480 AATCAGAAAGCTTCATGGCTTAAAGAGGAGTTGAATGCAATTTGTTCCGTTATTTGGAAGAA 539

DB 376 GAGATACTTCTGGAACTATCAAGAAACAGCTTGCAGCTTATAGCACCACTACTTGAACAG 435

QY 540 ATGAAAAAGAAAAGGTCGACGATGGACCACTTTGTTTCATGTCATAGCAGGATTAAG 599

DB 436 TTATGGCAACAAAAGGAAAGAAATTAAGGAGTTCTCGGATGTACAGTCACAGATCCAA 495

QY 600 AAAATTTGCTCTGAAATTAAGGCGCAGCG---ATTGTTTCCCTTTAAAGTTCGGTTGAT 656

DB 496 CAAATATGTGGAGAGATAGCTGGAACCTTGAACCTTAAATGATGTTTCACTGCGATTGAT 555

QY 657 CAGTCTGACCTCTCATTAAGAAAGCTTGAATGATGATGAGTGACGAAGACCTCGAATCCCTTCAG 716

DB 556 GAGTCTGACCTCTGCTGGAAGGTTGGATGAATATCAATCTGAGCTCCAAGAACTTCAA 615

QY 717 AAGGAGAGAGCGATCGGCTAAAGCAAGTATGATGATGAGTGACGAAGACCTCGAATCCCTTCA 776

DB 616 AAGGAGAGAGCGATGAGGTTGCACAAGGTTCTTGAATTTGTGAGTACTGTGATGATCTA 675

QY 777 TGTGAGTCTTGGCATAGATTTCAAGCAACAGTATATGAGTGACGCTGACCCCTAGCTTGGAC 836

DB 676 TGTGCTGCTTGGTATGAGCTTCTTCACTACTGCACTGAGGTTTCAATCCAGTCTAAAT 735

QY 837 GAAGCTGGAAG-----ATCAAGAAACCTGAGCAACACTTACAATTGAGAGGCTTGTGCT 890

DB 736 GACTCTACTGTTGTTTCAATCCCAAGAGCATAAAGTAAATGACACCCCTTCAAGGCTGGCTAAG 795

QY 891 GCGCAAAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGCTTCAAGATTTTGTCT 950

DB 796 ACCGCTTAAACGCTGAAAGAAATATAAAACAGAGGCTGCACAAAGTCCAAAGATTTAGCT 855

QY 951 TCTAGCATGCTGAGCTATGGAATCTCATGGATCTTCACTTGAAGAGCAGCAGATGTTTT 1010

DB 856 TCTCAGTTGATGATCTTTGGAAATCTAATGGATCTATCCCGGAGGAAAGGAGACTATTT 915

QY 1011 CAGAATATAACATGCAATATTTGCTGCTTCAGAACAAAGAGATAAATGAAACCAACCCCTC 1070

DB 916 GACCATGTTACCTGTAATATGTCAGCTTCTGTTGATGAAGTCACTGTCTCTGGTGCCCTT 975

QY 1071 TCACAGATTTCTGAAATTTATGTCGAATCTGAGGCTTTAAGGCTTGAACACTGAAAGCA 1130

DB 976 GCTCTGGATCTGATTCAGCAGGCTGAAAGTGGAGGACTTTCATCAGCTGAAAGCC 1035

QY 1131 AGTAAGATGAAGATCTTCTTTTAAAAAGAAAGCAGAACTAGAAAGCAGCAGTAAAGACGT 1190

DB 1036 AGCAGGATGAAGGAAATTTGCTTTTCAAGAAAGCAGAGAGCTCGAAGAGATATTTGCCCGT 1095

Qy	1191	GCTCATCTTGTGGCGAGGAAGTTATGCAAGAGAGTTTACATGTGAAGCTATTGTAAGCT	1250
Db	1096	GCTCATATGAAGTAGATCCAGATGCTGCCCGGAGAGAAGATTATGGCAITGATTGATTCA	1155
Qy	1251	GGAGCTATTGATCCCTCACCTAGTACTTTGAACAAATTTGAAGCTCACATTGCAACAGTCAA	1310
Db	1156	GGAAATATTGAACCAACTGAATTTACTGGCTGCATGGAACAATCAGATAGCAAAAGCAAAA	1215
Qy	1311	GAGGAAGCTTTTAGCCGGAAGGATATTCTTTGAGAAAGTTGAAAAGATGGCAAAATGCTTGT	1370
Db	1216	GAAGAAGCTTTAAGCCGAAAAGATATATTGGCAAGAGTTGAGAAATGGATGTCAGCATGT	1275
Qy	1371	GAGAGGAAGCCTGGCTGGGAAGATTACAAACAAAGATGATTAATCGTTACAATGTGGGAGG	1430
Db	1276	GAAGAAGAGAGTTGGCTTGAAGACTATAACCGGGATGAGAACAGGTATTAATGCAAGCAGA	1335
Qy	1431	GGAGCATCTAACACTTAAGAGGGCTTGAAAAGGCTCGTACTTTTGGTCAACAGATTCCT	1490
Db	1336	GGTGCAACACATAACCTCAACACGTGCAGAGAAAGACGGATATTGGTCAACAAAAATTCCA	1395
Qy	1491	GGAAATGGTAGATGTTTTTGAGAAACAAAAATTTGCTGCATGGAAAAATGAACGAGGAAAGGAG	1550
Db	1396	GCITTTGGTTGATACATTTGGTTGCTTAAACTCGTGCATGGGAAGAGATCATAGTA---TG	1452
Qy	1551	GATTTCAATATGATGGTGTGTAGCCTTTTCGTCAATGCTTGATGAATATATGTTTCGTTCGT	1610
Db	1453	TCATTTTACATATGATGGCGTTCTCTTCTTTGCCATGTTAGATGAATATGCCATGCTCAGA	1512
Qy	1611	CAGGAGAAAGACGACAGAGAGAGACAAAGGGATCAGAAAGATCCAGGATCAGCTC	1670
Db	1513	CATGAGCGGAAGAGGAAAAACGGAGGATGAGGGATCAGAAAAAGCATCACGAGCAGCGA	1572
Qy	1671	AAAGCGGAGCAGGAAGCTTTTGTACGGATCAAAACCCAGTCCATCCAGGCCCTTAAAGTACA	1730
Db	1573	AACACGGAACAAGAAACCATCTTTGGTTCAAGACCCAGCCCTGCTAGGCCAGTTAGTTCC	1632
Qy	1731	AAGNAGCCA	1739
Db	1633	AGTAAAGTCA	1641

RESULT 10

```

US-10-437-963-77259
; Sequence 77259, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77259
; LENGTH: 1680
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77174C.1
US-10-437-963-77259

```

Query Match 17.2%; Score 425.6; DB 17; Length 1680;  
Best Local Similarity 58.5%; Pred. No. 3.5e-109;  
Matches 856; Conservative 0; Mismatches 564; Indels 42; Gaps 5;

[illegible]

1089	CTCTGAGGCCAAAGATGAGGAGTTTCAGCAGGAAGGATGTTCTTCTGAGAAAGATGAGAAATG	1148
1358	GCAAAATGCTTGTGCAAGAGGAAGCCTGGCTTGGAGATTTACAAACAAAGATGATTAATCGTTA	1417
1149	GCAGGCTGCACTAGAGAGGGAATCTTGGCTTGAAGGAGTACAAACAGAAATGAGAAACAGATA	1208
1418	CAATGCTGGAGGGGAGCAATCTAACATCTAAAGAGGGCTGAAAGGCTCGTACTTTGGT	1477
1209	CAATGTGGGCAAAAGGGACTCATCTTGTGCTGAAGCGTGCAGAGAAAGCGCGTCCCTTGGT	1268
1478	CAACAAGATTCCTCGGAATGCTAGATGTTTTGAGNACAACAAAATTCCTGCATCGGAAAAATGA	1537
1269	CAGCAAAATGCCAGCAATGCGCAGAGGCCCTTGATCACAAGGTAATTCCTTGGGAGGAAGGA	1328
1538	ACGAGGAAACGAGGATTTTCATATGATGCTGTTAGGCTTTCGTCAATGCTTCATGAATA	1597
1329	GAGGGG---TGCCMAATTTGAGTACGATGGTGATGGTCTCTCGGACATGCTGGAGGAGTA	1385
1598	TATGTTCTGTTCTGTCAGGAGAAAGCAGNAGAGAGAAGAGAGACAAAGGGATCAGAAAGACT	1657
1386	CAACAATAACAAGGAAGGAGAAAGCAGCAGGAGAGAGAAAGAGACAAAGGGATCAGAGAAAGGAT	1445
1658	CCAGGATCAGCTCAAAAGCGGAG	1679
1446	GCTGGGTCAAGGACAGGGGAG	1467

RESULT 11

```

US-10-437-963-28884/c
; Sequence 28884, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 28884
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33439C.1
US-10-437-963-28884

```

	Query Match	15.8%	Score 389.8	DB 17	Length 2411	
	Best Local Similarity	55.9%	Pred. No. 6.3e-99			
	Matches 852	Conservative	0	Mismatches 647	Indels 24	Gaps 5
QY	240	ACATGCGATTCCGCTTCTACTCGAGCTCAATGTGTGATTTGGGATGAGTCTGCTGAGCCCGAC	299			
DB	2145	ACGTTGTGATCCTTTGCTGCAGAACTGCAGTTGATATGGGATGAAGTTGTTGAGAGCGAC	2086			
QY	300	ACGACGAGGGACAGGATGCTCTCGAGCTCGAGCAGGAGTCTCTGGAGGTCTTACAGCGG	359			
DB	2085	GAGGACCGCGACAAGGTCCTGTACCAGTTAGATCAGGAGTCCCTGGATGTTTACAAGG	2026			
QY	360	AAAGTTCACACGAGCGACCGGAGCCGCGCCAGCTCGGAGGCGCATTCGCGAGGGCGAG	419			
DB	2025	AAAGTTGACCAAGGCCACCGACTCTAGGATCTCCTTATCCAGGCGCTGGATGACTCGAAG	1966			
QY	420	GCAGAGCTCGCGCGGCATCTGCTCAGCCATGGCGAGCCGCCGCTGCACGTAGACGTCA	479			
DB	1965	ATTGAGCTGGCCAGACTTTTATCTGCTCTAGGACGAGGCGCAT---AGCAAGAACTCTCT	1909			



QY 1545 AAGGAGGATTTTACATATGATGGTGTAGCTTTTGGTCAATGCTTGTGATGAATATATGTTTC 1604  
Db 829 --TCTGTCTTTATGTATGATGGTGTACCTCTTCTAGCTATGTTGGATGAATATGTCATG 772  
QY 1605 GTTCGTAGGAGAAAGAGCAAGAGAGAGACAAAGGATCAGAAG---AAGCTCCAG 1661  
Db 771 CTAAGGCGAGAAAGGGAAGAGATGAAGAAAGAAATGAGGGAACAAAGCGCTACATCGAG 712  
QY 1662 GATCAGCTCAAGCGGAGCAGGAAGCTTTGTACGGATCAAAACCCAGTCCATCCAGGCC 1721  
Db 711 CAACAACTAGAACACCGACCATGAAGGCCATTTGGTTTACGTGTGAGCCCTAACAGACCA 652  
QY 1722 CTAAGTACAAAGAGCCACTAG 1744  
Db 651 GTCAGTGCAAGAGGATACCTCG 629

RESULT 12  
US-10-424-599-111490/c  
; Sequence 111490, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 111490  
; LENGTH: 2712  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_71686C.1  
US-10-424-599-111490

Query Match 15.8%; Score 389.6; DB 16; Length 2712;  
Best Local Similarity 55.0%; Pred. No. 7.8e-99;  
Matches 842; Conservative 0; Mismatches 669; Indels 21; Gaps 3;  
QY 238 CGACATGCGATTCGCTTCTACTGGAGCTCAATGTGATTTGGATGAGGTGCGGTGAGCCG 297  
Db 2241 CTACTTGCACTCTTGTCTCAGAGAACCTTGAGCAATATGGAACATATTTGGGAGACCG 2182  
QY 298 ACACAGCAGGAGGACAGGATGCTGCTGGAGCTCGAGCAGAGTGCTTGGAGGTCTACAGGC 357  
Db 2181 AAGTGGAGAAAGATCGAATGCTGATGGAGCTGGAGAGGGAATGCTTGAAGATATACAGGA 2122  
QY 358 GGAAGTGCACAGCGGAACCGAGCCGCGCCAGCTCGGAGAGCCATCGCGAGGGCG 417  
Db 2121 GAAAGGTTGATGAGGCTGCTAAACCAAGCACGCTTTCATCAACAGTTGCGCGCAAG 2062  
QY 418 AGGAGAGCTCGCCGCACTGCTCAGCCATGCGGAGCGCCCGCTGCA---CGTTAGAC 474  
Db 2061 AMGCAGAGCTTGCTACATGATGGCTGCACCTTGGAGAACATGATATTCATTCACCGATTA 2002  
QY 475 AGTCAAAATCAGAAAGCTTTCATGGCTTAAGAGAGGAGTTGAAATGCAATTTGCTCGTATTTGG 534  
Db 2001 AGATGGAGAGAGATCAGATCCCTGGAAGCAGAAATTTGCGTCCATCACACCTTTGGTTG 1942  
QY 535 AAGAAATGAAAGAAAGAAAGTGCAGCAATGGAACAGCTTTGTTTCATGTCATAGAGCAGA 594  
Db 1941 AAGAACTGAAGAAAGAAAGAAAGATGAAGATTTGAAGCAATTTGAAGATGTAAGAGACTCAA 1882  
QY 595 TTAAGAAATTTGCTGCTCAATTAAGCCAGCCGATTTTGTTC-----CCTTTA 642  
Db 1881 TAGAAAGATTAAGTGGAGAGATTTTGGATTTCCATCTGTCAATTAATGCTTTGAGCAGCA 1822  
QY 643 AAGTTCGGTTGATCAGTCTGACCTGTCTAATTAAGAAAGCTTTGATGAGTTGACGAAGGACC 702

Db 1821 CAGCAGTTGAGGATGAACAGGACTTGTCACTTAGGAGACTTAATGAATATCAAAACGCATC 1762  
QY 703 TGGAAATCCCTTCAGAAAGGAGAGCCGATCGGCTAAAGCAAGTATAGAAACATTTGAAT 762  
Db 1761 TCCGAACACATTTCAAAAAGAAAAGTCTGATCCCTTCAAAAAGGTCTTCAATTCGCGTGAATG 1702  
QY 763 CTTTTCATTCCTTATGTGAGGTGCTTGGCATAGATTTCAAGCAAAACAGTATATGAGGTGC 822  
Db 1701 AGTTCATCTCTTTGTAGTGTGCTTGGTGGATTTTGGTCAGCCGCGGTGATGTAC 1642  
QY 823 ACCCTAGCTTGGACGAAGCT-----GAAAGGATCAAAAGAACCTGAGCAACACTACAAT 876  
Db 1641 ATCCAAGTTTGCATGGGACTCAGGTGGAACAGTCAACTAATATTAGCAATAGCACATTTGG 1582  
QY 877 AGAGGCTTGTGCTGCCGCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGC 936  
Db 1581 AAGTCTAGACGAGGCCATTTCTCAAGTTTAAAAATAGAAAGGAAAACCTAGAAATTCAGAA 1522  
QY 937 TTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGATACTCCACTTTGAAG 996  
Db 1521 TGAAGGATTTGTATCTAAACTATTTGAACTTTGGAATTTGATGACTCATCAAAAGAG 1462  
QY 997 AGCAGCAGATGTTTTCAGAAATATAACATGCAATATTGCTGCTTCAGAAACAGAGATACTG 1056  
Db 1461 AGAGAACTGTTTATGAAGATTAATCTTCTTATTTGGAACCTCAGAAATCAGAAATCACTG 1402  
QY 1057 AACCAAAACACCTCTCCACAGATTTCTCTGAATATTGCAATCTGAGGTGTTAGGCTTTG 1116  
Db 1401 AACGAGTGTCTTTTCAACAGAAATGATAGAAAGGCTTCGGCAGAAATGTCAGAGGCTTTG 1342  
QY 1117 AACAACTGAAAGCAAGTAAGATGAAGATCTTTGTTTAAAAAAGAAAGACAGAACTAGAA 1176  
Db 1341 CCAAACTTAAAGCAAGCAGATGAAGAACTTTGTTTAAAGAGGCTCAGAGTTAGAGG 1282  
QY 1177 AGCATAGAAGACGTGCTCATCTTGTTCGCGAGGAGGTATGTCAGAGAGGTTTAGCATTTG 1236  
Db 1281 AAATATGTAAATTTGACTCATACTGAACAGATACAAGTACTACAGCCGAGAAAGCTAGTG 1222  
QY 1237 AAGCTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTTGAACAAATTTGAAGCTCACA 1296  
Db 1221 CATTAATAGATTTCTGGAGCTGGTGGATCTTCTGAGCTATTAGCCAACTATTGAAGCAGAGA 1162  
QY 1297 TTGCAACAGCTGAAAGAGAGAGCTTTTAGCCGGAAGGATATTCTTTGAGAAAGCTTCAAAGAT 1356  
Db 1161 TAATAAGGCAAAAGATGAAGCTTTGAGCAGAAAGAAAGTAACAGATAGGATTTGATAGT 1102  
QY 1357 GGCAAAATGCTTTGAAAGAGAGAGCCCTGGCTGGGAAGATTACAAACAAAGATGATATCGTT 1416  
Db 1101 GGTTCGCTGCTGTGAAGAGAGAGAAATTTGGCTTGAAGAAATATAATCAAGATGACAAAGAT 1042  
QY 1417 ACATGCTGGGAGGAGCAGCTCACTTAACATTAAGAGGCTTGAAGAGGCTGTAAGGCTGTTGG 1476  
Db 1041 ACTGCTGGGAGGAGGCTCAGCTTAATCTTAAACGTGAGAAACGCGCTAGAAATAACTA 982  
QY 1477 TCAACAAAGATTTCTTGAATGTTAGATGTTTGAAGCAAAAAATTTGCTGATGGAATAATG 1536  
Db 981 TAAGCAAAATTTCCAGCTATTGTTGACAACTTTATAACAAACAACTAGACCTGGGAATCGG 922  
QY 1537 AACGAGAAAGGAGGATTTTCAATATGATGGTGTGTTAGCTTTTCGTAATGCTGATGAAT 1596  
Db 921 CTGAAAAAGAGACTCAATTTCTATATGATGGGTACGCTTTGGTGTCAATTTTGGACGAT 862  
QY 1597 ATATGTTTCGTTCTGAGGAGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1656  
Db 861 ATAACTGGCAGACAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 802  
QY 1657 TCCAGGATCAGCTCAAAAGCGGAGCAGGAAGCTTTTGTACGGATCAAAACCCAGTCCATCCA 1716  
Db 801 TGCAGATCTACTCTCTTAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 742  
QY 1717 AGCCCCCTAAGTACAAAGAGGCGACCTAGGCAC 1748



Db 741 AAAATAACAGCTTTTAGAAAGACTAATAGTTAC 710

RESULT 13

US-10-425-115-39061  
 ; Sequence 39061, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 39061  
 ; LENGTH: 2422  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_135624C.1  
 US-10-425-115-39061

Query Match 15.4%; Score 379; DB 18; Length 2422;  
 Best Local Similarity 55.3%; Pred. No. 7.3e-96;  
 Matches 830; Conservative 0; Mismatches 650; Indels 21; Gaps 4;  
 QY 237 ACGACATGCGATTGCGCTTCTACTGGAGCTCAATGTGATTGGATGAGTCGGTGAGCCC 296  
 Db 552 ACGACATGCGGCTCTTGTCTGAGCACTTCTAGCTGATGATGGATGAGTTCGCGAAGT 611  
 QY 297 GACACGACGAGGACAGGATGCTGCTGGAGCTCGACGAGGAGTCTGAGGTCTACAGG 356  
 Db 612 GATGAAGATCGTGACAAGATTGTTACTTTCAGCTAGAGCAAGAGTCTTAGATGTTATAGG 671  
 QY 357 CGGAAGTGCACGAGCGAAGCGGAGCCGCGCCAGCTCGGAGGCCATCCCGAGGCC 416  
 Db 672 AGAAGGTTGATCAGGCTTCTAGCTCAGGTCTATGCTCTCCACACAGCTTTCCAACTCC 731  
 QY 417 GAGGACAGCTCGCCGGCATCTGCTCAGCATGGGCGAGCCGCGCTGACGCTTAGACAG 476  
 Db 732 AATCGGAAGTCTGATGATCTCTTCTGATGAGGAGTCTATTTCTGCTGCTC--- 789  
 QY 477 TCAATCAGAAGCTTCATGGCTTAAGAGAGGAGTGAATGCAATGCTTCGGTATTTGGAA 536  
 Db 790 -CTGACAAGACATCTGGTACAACTCAAGGAGCACTAGCAGCTATATCACCATCTTTAGAG 848  
 QY 537 GAAATGAAAAGAAAAGGTGCAACGATGCAACGAGTCTGATGATGATGATGATGATGATG 596  
 Db 849 ACACCTCCGAGGAAAAGAGAGTAGGGTGAAGAGTGTGCTCATGTACAACCTTCAGATT 908  
 QY 597 AAGAAAATTCGCTCAAAATTAAGCCAGCGGATTTGTTCCCTTTAAAGT-----T 647  
 Db 909 CAGATCTCTGATGAAATTCGCTGGGATCTACATATTTGGTGAACAGTTGGGAACACCC 968  
 QY 648 CCGGTTGATCAGTCTGACCTGTCTAATGAAGAGCTTGTGATGATGATGATGATGATGATG 707  
 Db 969 CATGTTAATGCGGATGATCTTTTCAGTGAGGAGATTAAATGAATATCTTTCTGAACATCAA 1028  
 QY 708 TCCCTTCAGAGAGAGAGCGATTCGGCTTAACCAAGTGTAGAAATTTGAATTTCTTTG 767  
 Db 1029 GCCCTGCAAAAGAAAAGAGTACTAGGCTCTCAAGGATTTCTGAATCCGAGTTCAGTG 1088  
 QY 768 CATTCTTATGTGAGGTGCTTGGCATAGATTTTCAAGCAACAGTATATGAGGTGACCCCT 827  
 Db 1089 CAGATCTCTGTTCTGTTACTTGGCATGATTTTGTGGGCACTGTTACTAAAGTTTCATCCA 1148  
 QY 828 AGCTTGACCAAGCTGAAG-----ATCAAAGAACCTGAGCAACACTACAATTTGAGAG 891  
 Db 1149 AGTCTTGATGACTCTGTTGGTGTCCCAATCCAGAGGATTTAGTGGCGAAACAATATCCAG 1208

QY 882 CTTGCTGCTGCCCAAAACAGACTCGCTGAAATGAAGATCCAAAGATGCCAAAAGCTTCAA 941  
 Db 1209 CTGCTTAAATGGTGTGTTGTTACTTTGAAGAGAAAATCGAAGAGGTTTGCAAAAGATTCAA 1268  
 QY 942 GATTTGCTCTAGCATGCTGAGCTATGGATCTCATGGATATCTCCACTTTGAGAGCAG 1001  
 Db 1269 GCTCTAGCTTCCAGCTATCTGATCTTTGGAACTTTAATGGATTTCTCTGTGGAGAACAG 1328  
 QY 1002 CAGATGTTTTCAGAAATATAACATGCAATATCTGCTTCAGAAACAAGAGATAACTGAACCA 1061  
 Db 1329 CAACCTTTTCTGCTAGTAAACATGATATGTTCTCAACATTTGGATGATGTGATAGTTCCA 1388  
 QY 1062 AACACCTCTCTCCACAGATTTTCTGAAATATGCTGCAATCTGAGGTGTTAAGGCTTGAACNA 1121  
 Db 1389 GGAGCTCTGGCTCTTGTATGTAATTTCAGCAGCGGAACTCGAAGTTTGAAGGCTTGTATCAG 1448  
 QY 1122 CTGAAGCAAGTAAAGTAAAGTCTTTGTTTAAAGAAAGAGCAGAACTAGAACGAT 1181  
 Db 1449 CTAAGAGCTAGTAGGATGAAGGATATTGCAATTCAGAAACACAGCTGAACCTTGAAGATATA 1508  
 QY 1182 AGAAGACGTCTCATCTTTGTTGGCAGGAAGGTTATGCAAGAGGAGTTTAGCATTTGAAGCT 1241  
 Db 1509 TATGCCCGGCTCATGTTGCTAGTAGATTCTAGTCTGCAAGAGATAGGATATGCTATT 1568  
 QY 1242 ATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTTGAAACAAATTTGAAGCTCACATTGCA 1301  
 Db 1569 ATTGAGTCCAGTAGTTTCGAGCTTTCAGAACTTACTGCTGATATGGAAGAACCCAGATCTT 1628  
 QY 1302 ACAGTGAAGAGGAGGCTTTTACCGGAAGGATATTCTTTGAAAGATTTGAAAGATGGCAA 1361  
 Db 1629 AAAGCAATAGAGGATCTCTTAAGCAGAAAGGACATATTGGAGAGGGTTGATAGGTGGATG 1688  
 QY 1362 AATGCTTGAAGAGGAGGAGCTGCTGGAAGATTTACAAACAAGATGATAATCGTTACAAT 1421  
 Db 1689 TCAGCATGTGAAGAAAGAAAGCTGGCTTGAAGCTACAGCCGGAAGATTAACAGGTACACT 1748  
 QY 1422 GCTGGAGGAGGAGCAGCATCTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTTGGTCAAC 1481  
 Db 1749 GCAACTCGGGGTGCAACATCTGAATCTGAAGCGTGCAGAAAGGCGCTGTTTGGTCAAT 1808  
 QY 1482 AAGATTTCTGGAATGGTAGATGTTTGGAAACAAAATTTGTCATGGAAGAAATGAACGA 1541  
 Db 1809 AAAATTCGGCTATTGTTTGACACACTGTTGCAAAAGACCCGGGATGGAAGAACAGAGCAT 1868  
 QY 1542 GGAAGAGGAGGATTTCACATATGATGCTGTAGCTTTTGTCAATGCTTGTGATGAATATATG 1601  
 Db 1869 GG---CGTGGCTTCACCTACGACGGCTCCCTCTTTCTTGGCTGCTGACGAATACAAA 1925  
 QY 1602 TTCGTTCTGAGGAGAAAGAGCAAGAGAGAGAGCAAGAGGATCAGAAAGAGCTCCAG 1661  
 Db 1926 ATCTTGAGCAGGAG 1985  
 QY 1662 GATCAGCTCAAGCGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1721  
 Db 1986 GAACAGCTAGCCGCGCAACAG 2045  
 QY 1722 C 1722  
 Db 2046 C 2046

RESULT 14  
 US-10-425-114-35072  
 ; Sequence 35072, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 35072

; LENGTH: 1788

; TYPE: DNA

; ORGANISM: Zea mays subsp. mexicana

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLTEBSINTE071C04\_FLI

US-10-425-114-35072

Query Match 15.3%; Score 377.6; DB 16; Length 1788;

Best Local Similarity 55.4%; Pred. No. 1.5e-95;

Matches 825; Conservative 0; Mismatches 644; Indels 21; Gaps 4;

QY 237 ACGACATGCGATTCTGCTTCTACTGAGCTCAATGTGATTGGATGAGTGGGTGAGGCC 296

DB 302 ACGACATGCGTCTTCTGCTGAGCACTTCAGCTGATATGGATGAGGTGGCGAAGT 361

QY 297 GACACGACGAGGACGAGATGCTGTGAGCTCGAGAGGAGTGCCTGAGGTCTACAGG 356

DB 362 GATGAAGATCGTGACAAGATGTACTTCAGCTAGAGCAAGAGTGTCTAGATGTTATAGG 421

QY 357 CGGAAGGTGACGAGCGAAGCGGAGCGCGCCAGCTGCGGAAGGCCATCGCCGAGGC 416

DB 422 AGAAGGTGATCAGGCTTCTAGCTCCAGGTGATCTCTCTCCCAACAGCTTCCCACTCC 481

QY 417 GAGGACAGAGCTCGCGGATCTGCTCAGCCATGGCGGAGCGCCGCTGACGTTAGACAG 476

DB 482 AAATCGGAACAGCTAGCTCTCTTCTGATAGGAGTGCATCTATTTCTGCTGCTC-- 539

QY 477 TCAATATCAGAGCTTCATGGCTTAAGAGAGGAGTTGAATGCAATGTTTCCGTATTTGAA 536

DB 540 -CTGACAAGACATCTGGTACAATCAAGAGCAACTAGCAGCTATATCAACCATCTTTAG 598

QY 537 GAAATGAAAAGAAAAGGTGGAACGATGGAAACAGTTTGTTCATGTCATAGACAGATT 596

DB 599 ACACCTCGGAGGAAAAGAGATAGGTGAAAGTGTGCTCATGTACAACCTTCAGATT 658

QY 597 AAGAAAATTTCTGCTGAATAAGGCGAGCGGATTTGTTTCCCTTT-----AAGTT 647

DB 659 CAGATACCTCGTGATGAAATGCTGGGATCTACATATTTGTGAACAGTTGGAACACCC 718

QY 648 CCGGTTGATCAGCTGATCTGATCAATTAAGAACTTGATGAGTTGACGAAGCACTGGAA 707

DB 719 CATGTTAATCGGATGATCTTTTCAGTGAGGAGATTAAATGAATATCTTTCTGAACACAA 778

QY 708 TCCCTTCAGAGGAGAGGAGCGGCTAAGCAAGTGATAGCAATTTGCAATTTCTTTG 767

DB 779 GCCCTGCAAGGAAAAGAGTACTAGGCTCTCAGGATCTTGAATCCGTGAGTTGAGT 838

QY 768 CATTCCTTATGAGGCTGTTGGCATAGATTTCAGCAAAACAGTATATGAGGTGCACTCC 827

DB 839 CACGATCTCTGTTGTTGCTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 898

QY 828 AGCTTGGAGAGCTGAAAG-----ATCAAGAACTGAGCAACACTACAAATTTGAGAGG 881

DB 899 AGCTTGTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 958

QY 882 CTTCGCTGCGGCAACAGAGCTGCGTGAATGAAGATCCAAAGGATGCAAAAGCTTCAA 941

DB 959 CTGCTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1018

QY 942 GATTTTGTCTTACGATCTGAGCTATGGAATCTCATGGATCTTCCATTTGAAAGAGCAG 1001

DB 1019 GCTCTAGCTTCCAGCTATCTGATCTTTGGAACCTTAATGGATTTCTCTGTGGAGGAA 1078

QY 1002 CAGATGTTTCAAGATATACATGCAATATTCCTGCTTCAGCAACAGAGATTAACCTGA 1061

DB 1079 CAACCTTTTCTGTCAGTAACTGCAATAATGCTCCTCAACATTTGGATGATGTTGATGTTCCA 1138

QY 1062 AACCCCTCTCCACAGATTCTTCTGAATATGCTGAATCTGAGTGTAAAGCTTGAACAA 1121

DB 1139 GGAGCTCTGGCTCTTGTATGTAATTCAGCAGCGGAATCTCGAAGCTTGAAGCTTGTATCAG 1198

QY 1122 CTGAAGCAAGTAAGATGAAGATCTTGTGTTTAAAAAAGAACAGCAACTAGAAGAGCAT 1181

DB 1199 CTAAAGCTAGTAGGATGAAGGATATTGCATTTCAAGAAACAGACTGAACCTTGAAGATATA 1258

QY 1182 AGAAGAGCTGCTCATCTTGTGTCGAGGAAGGTTATGCAAGAGGAGTTTAGCAATTTGAAGCT 1241

DB 1259 TATCCCGGGCTCATGTTGTCAGTAGATTCTAGTCTCTCAAGAGATAGGATAATGCTATT 1318

QY 1242 ATTGAAGCTGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTTCACATTGCA 1301

DB 1319 ATTGAGTCCAGTAGTTTCGAGCCTTCAGAACTCTGCTGATATGGAGAACCAAGATATT 1378

QY 1302 ACAGTGAAGAGGAGGCTTTTAGCCGGAAGGATATTCTTGAGAAAGTTTGAAGATGGCAA 1361

DB 1379 AAAGCAATGAGGAGTCTTTAAGCAGAAAGGACATATTGGAGAGGTTGATAGTGGATG 1438

QY 1362 AATGCTTGTGAAGAGGAGGCTGCTGGAAGATTACAAAGATGATATATCGTTTACAAT 1421

DB 1439 TCAGCATGTGAAGAGGAGGCTGCTTGAAGACTTACAGCCGGGACGATAACAGGTACAGT 1498

QY 1422 GCTGGAGGGGAGCACATCTAAACACTAAGAGGCTCGAAAGGCTCGTACTTTTGGTCAAC 1481

DB 1499 GCAACTCGGGGTGCACATCTGAATCTGAAGGTCGAGAAAGGCCGCTGTTTGGTCAAT 1558

QY 1482 AAGATTCCTGGAATGTTAGTATGTTTGGAGAACAAAAATTTGTCATGGAATAATGAACGA 1541

DB 1559 AAATTCGGCTATTGTTGACACACTCTTTGCAAGACCCCGGCGATGGGAACAGAGCAT 1618

QY 1542 GGAAGAGAGATTTCACATATGATGTTGAGCTTTGAGCTTTGCTCAATGCTTGAATATATG 1601

DB 1619 GG---CGTGGCTTCACTACGAGCGCTCCCTCTTCTGCGATGCTGGAACGAATACAAA 1675

QY 1602 TTGCTTCTGAGGAGGAGGAGCAAGAGAGAGCAAGAGGATGAGGAGGATCAGAAAGAGCTCCAG 1661

DB 1676 ATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1735

QY 1662 GATCAGCTCAAGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1711

DB 1736 GAACAGCTAGCCGCGCAACAGGAGAGCTGTTGCGGTCAAAACCGAGCCC 1785

#### RESULT 15

US-10-437-963-32793

; Sequence 32793, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; SEQ ID NO 32793

; LENGTH: 2345

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_36968C.1

US-10-437-963-32793

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 32793

; LENGTH: 2345

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_36968C.1

US-10-437-963-32793

Query Match		14.0%;	Score 346.4;	DB 17;	Length 2345;
Best Local Similarity		54.1%;	Pred. No. 1.2e-86;		
Matches 826;		Conservative 0;	Mismatches 676;	Indels 24;	Gaps 5;
QY	237	ACGACATGCGATTGCTTCTACTGAGAGCTCAATGTGATTTGGATGAGTTCGGTGAGCCC	296		
DB	373	ACTTCGTGTGGATCATTTGCTGCAAAACTCGCAGTCGGTATGGATGAAGTTGGTGAAGC	432		
QY	297	GACACGACGAGGACGAGTGTCTGGAGCTCGACGAGGAGTGTCTGGAGTCTACAGG	356		
DB	433	GAGGAGGACCGGACAAAGTCTTATTTGAGCTGGACGAGGATTTGGATGCTACAG	492		
QY	357	CGGAAGGTCCGACGAGCGAACCGGAGCCGCGCCAGCTCGGGAAGGCCATCCGCGAGGCG	416		
DB	493	CGGAAGTTGATCAGCAACCAATCAAGGGATCTTCTCCAGGCACTGGATTACTCA	552		
QY	417	GAGGACAGACTCGCGGCATCTGCTCAGCCATGGGCGAGCCGCGTGCACGTTAGACAG	476		
DB	553	AAGACAGAGCTTGCTAGGCTTGCTTCTGCGCTTGGTGAATAATCCATAGATATAAGTC--	610		
QY	477	TCNAATCAGAAGCTTCATGCTTTAGAGAGGAGTTGAATGCAATTCGTCGATTTGGAA	536		
DB	611	-CTGAGAAACAGCAGCGACTATCAAGGAACAACTAACTGCTATAGCTCCAACTTGA	569		
QY	537	GAAATGAAAAGAAAAGGTGCAACCGATGGAACCAAGTTTGTTCATGCTATAGAGCAGATT	596		
DB	670	CAATTGGGCAAGAGAAAAGGAGAGATAAAGGATTTAGCTAAATATACAGTCAAGAA	729		
QY	597	AAGAAAATTCGCTCGAATAAAGGCCAGCGGATTTTGTTCCTTTAAAGTTCC-----	649		
DB	730	GAGCAATACGTTGGTGAGATTGCTGTAATCTGAGATGGGACAGCAAGTGGCACTACCA	789		
QY	650	--GGTTGATCAGTCTGACCTGCTATTAGAAAGCTTGATGAGTTGACGAAGCACTGGAA	707		
DB	790	CAGATTAATGAGGATGATTTGACAGTTTCGGAAGCTTCGAGAGTTCCAGTTACAGCTTCAA	849		
QY	708	TCCTCTCAGAAGGAGAGAGCGATCGCTAAAGCAAGTGATAGAACTTTGAATTTCTTTG	767		
DB	850	GAGCTTGAGAGAGAGAGTGGTCTGAGCTGGAAAGGTTCTTGAGCATGTTGATGTTGA	909		
QY	768	CAATTCCTTATGTAGGTGCTTGGCATAGATTTCAAGCAAAACAGTATATAGGTGCAACCT	827		
DB	910	CATGATCTATGCAATGCTGCTGGATGGATTTTATGAACAATAACTCAAGTCCATAGC	969		
QY	828	AGCTTGGACAGACTGAAGS-----ATCAAGAACTGACCAACACTCAATTTGAGAGG	881		
DB	970	AGTCTAGATGACTCTATTGGTAAACGAACCAAGCAATTAAGCAATGAAACCCCTGTCAAAA	1029		
QY	882	CTTGCTGCTGCCGCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGCTTCAA	941		
DB	1030	CTTGATAGGACCATTGGTACTCTTATAGGACAAAGGTTGAGACTTGGAAGCTTCAA	1089		
QY	942	GATTTTCTTCTAGCATGCTCGAGCTATGGAATCTCATGGATATCTCCACTTGAAGAGCAG	1001		
DB	1090	GAGCTTGCCACTCAACTCTATGATCTATGGATCTTATGGACACCCCGTGGGAAGAAAG	1149		
QY	1002	CAGATGTTTCAGAAATATAACATGCAATATGCTGCTTCAGAAACAGAGATACTGAACCA	1061		
DB	1150	AGCTTTGTTGATCATGTTCTCGTCAATAGAACAGCAACTGTTGAGGAGGTCTGTTTCCT	1209		
QY	1062	AACACCTCTCCACAGATTTCTCGAATATGTCGAATCTGAGGTGTTAAGGCTTGAACAA	1121		
DB	1210	GGTGCACTTCTGTTGATGTTGATGATCAAGCACAGACTGAGGTTGNAAGTTGGATCAG	1269		
QY	1122	CTGAAGCAAGTAAAGATCTTTGTTTAAAGAAAGCAAGCAACTAGAAAGCAT	1181		
DB	1270	CTAATAACAGCAAGATGAAGAAATAGCTTTTAAAGAAAGCAAGCACTCTGGAAGATATT	1329		
QY	1182	AGAAGACGCTCATCTTTGTTGGGAGGAAGGTTATGCAGAGGAGTTTACATTGAGCT	1241		
DB	1330	TATGCTAGCACTCATGTTGTTATTAGACACAGCAGTTGCCCATGAGAAATACAGGCAC	1389		
QY	1242	ATTGNAAGCTGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCACATTGCA	1301		

Search completed: November 20, 2004, 00:37:52  
Job time : 1219 secs

DB	1390	ATTGAGTCAGGGAACATGGAACCTTTCAAGACTGATTTGTCATGGATAGTTCAGATACTG	1449
QY	1302	ACAGTGAAGAGGAGGCTTTTAGCCGAAGGATATTTCTTGAGAAAGTTTGAAGATGGCAA	1361
DB	1450	AAAGCAAGAGGAGGAGCTCTGAGCAGAAAGAAATATTAGATAAAGTTGAGAGATGGATA	1509
QY	1362	AATGTTGTGAAGAGGAAAGCCCTGGCTGGAAGATTACAACAAAGATGATAATCGTTACAT	1421
DB	1510	TCCTCATCGAGGAGGAAAGCTGGCTTCAAGACTATAGCAGAGATGACAAACAGGTATAAC	1569
QY	1422	GCTGGAGGAGGACACATCTAACACTAAGAGGCTGAAAGGCTCGTACTTTTGGTCAAC	1481
DB	1570	TCAGGCCAGGTGCTCACCTGAATCTCAACGTCGAGAAAGGCCGCTATTCTAGTCAGC	1629
QY	1482	AAGATTCTGGAATGCTAGATGTTTTGAGAAACAAAAATTTGCTGCAATGGAATAATGAACGA	1541
DB	1630	AAGATTCAGCCCTTGTGAAACTCTGGTGGCTAAGACGAGAGCTTGGGAAGAAAAATCAC	1689
QY	1542	GGAAAGGAGGATTTTCAATATGATGTTTACGCTTTCGTCATGCTTGTGATGAATATATG	1601
DB	1690	GG---GCTGCCATTTATGATGACCGTGTCTCTGCTAGCAATGCTGGATGAGTATGTC	1746
QY	1602	TTGCTTCTCAGGAGAAAGCAAGAGAGAGACAAAGGGATCAGAAAGAGCTCCAG	1661
DB	1747	ATTCTTAGGCAAGAGAGGGAAGAGAGAAAGAAATGCGGGAAACAAAACGCCAGACG	1806
QY	1662	GATCAG---CTCAAAGCGAGCAGGAAGCTTTGTACGGATCAAAACCCAGTCCATCCAAG	1718
DB	1807	GAACAGCTACTTAACATCGACCGTGGGGCCATTCGGNACACGGGTCAATCCCTACAGA	1866
QY	1719	CCCCTTAAGTACAAAGAAGGCACCTAG	1744
DB	1867	GTGACCAGTGCACAAAGAGTGGCTGG	1892

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:08:20 ; Search time 101 Seconds

(without alignments)  
3930.777 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKQDLHQMSCTTCDLSL.....DIDYSFERRLAIFYLARQWV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:\*

1: uniprot\_prot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1929.5	55.2	730	2	Q6L4A9 solanum dem
2	1299.5	55.2	730	2	Aat40494 solanum d
3	1900.5	54.4	707	2	Q9FHM4 arabidopsis
4	1625.5	46.5	645	2	Q8LQE3
5	1362.5	39.0	648	2	Q9L2Y0
6	1314.5	37.6	587	2	Q9FLP0
7	1311	37.5	580	2	Q9FEV9
8	1296	37.1	576	2	Q84VU1
9	1279.5	36.6	584	2	Q9FEV8
10	1268.5	36.3	582	2	Q9FEV7
11	1236.5	35.4	578	2	Q8LQEG
12	1208	34.6	608	2	Q9SIS3
13	1203.5	34.4	603	2	Q8LQ36
14	1202	34.4	511	2	Q9LVB1
15	1182.5	33.8	589	2	Q6Z113
16	1182.5	33.8	589	2	BAD12872
17	1181.5	33.8	559	2	Q6Z2N0
18	1181.5	33.8	559	2	BAD08014
19	1168.5	33.4	635	2	Q9LQW1
20	1127.5	32.3	607	2	Q8H006
21	1066	30.5	570	2	Q61569
22	1042	29.8	587	2	Q9ZVJ3
23	1004	28.7	473	2	Q9S216
24	966	27.6	626	2	Q852G8
25	905	25.9	592	2	Q9C7G0
26	878.5	25.1	488	2	Q6Z166
27	878.5	25.1	488	2	BAD09028
28	354	10.1	609	2	Q7T0Z3
29	334	9.6	605	2	Q802Y9
30	334	9.6	605	2	Q6D126
31	331	9.5	603	2	Q99K43

32 323.5 9.3 620 2 043663 043663 homo sapien  
33 307 8.8 516 2 08CE25 08CE25 mus musculus  
34 294.5 8.4 581 2 07ZXD0 07ZXD0 xenopus lae  
35 280.5 8.0 576 2 09BSB6 09BSB6 homo sapien  
36 275.5 7.9 598 2 06DI02 06DI02 brachydanio  
37 275.5 7.9 845 2 06FKN0 06FKN0 candida gla  
38 271.5 7.8 801 2 06CJS4 06CJS4 kluyveromyc  
39 267.5 7.7 407 2 06LS12 06LS12 oryza sativ  
40 267.5 7.7 407 2 AAT39217 AAT39217 oryza sat  
41 264.5 7.6 595 2 07ZVK0 07ZVK0 brachydanio  
42 262 7.5 579 2 06DI23 06DI23 brachydanio  
43 259.5 7.4 885 1 ASE1 YEAST ASE1 YEAST  
44 254.5 7.3 888 2 Q7SE4 Q7SE4 ashbya gos  
45 254.5 7.3 888 2 AAS50503 AAS50503 ashbya go

## ALIGNMENTS

### RESULT 1

Q6L4A9  
ID Q6L4A9 PRELIMINARY; PRT; 730 AA.  
AC Q6L4A9;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Putative microtubule-associated protein.  
GN Name=PGEC561.15;  
OS Solanum demissum (Wild potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=50514;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Buell R.C., Liu J., Zaborisky J., Tallon L., Baker B.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ronning C.M.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Buell R.;  
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: AC136471; AAT40494.1; -;  
DR InterPro: IPR007145; MAP65\_ASE1.  
DR Pfam: PF03999; MAP65\_ASE1; 1.  
SQ SEQUENCE 730 AA; 82876 MW; E4E80EBD8F60D4E5 CRC64;

Query Match 55.2%; Score 1929.5; DB 2; Length 730;  
Best Local Similarity 57.1%; Pred. No. 1.5e-79;  
Matches 418; Conservative 91; Mismatches 154; Indels 69; Gaps 15;

QY 12 MSTTCDLSLLELVINWDEVEGPTTDRMLLEBCECLVYRKVDQANRSRAQLKAIA 71  
DB 1 METTCGLLHLELIQIINWDEVEGCTDRMLPBIERECLVYRKVDQANRSRAQLKAIA 60  
QY 72 EGAEALAGTCSANGPEPPVVRQSKLHLGRLELNAIVPYLEMKKKVYVHVVIE 131  
DB 61 DSEAEALATCSANGPEPPVHIKQSKDSQGLKAEALRAVLPELBEMRKSRDKNQFIEVMK 120  
QY 132 QIKTSSSEIRPADFPVPKVPVQSDLSLRKLDLT KDLESLOKESKDRKQVTEHLSLH 191  
DB 121 QITKINEI--YRFTSASLVDESLSLRKLELTELTLQKEKERLQVNLHGLTN 178  
QY 192 SLCEVLGIDFKQTVVEVHPSPDEAGSKNLSNTTIERLAAAANRLRMIQRMKQLDFA 251  
DB 179 SLCSVLGMDFKHTINEVDNLSGESEAKNICDDTIQNLAAITQRLQEVKLQRMQLDLT 238  
QY 252 SSMLELWNLMDTFLPEEQMFQNTCNIAASEQIEPTNTLSTDFLNVYSEVLRSLQKA 311  
DB 239 TSMLELWNLMDTFLPEEQMFQNTCNIAASEQIEPTNTLSTDFLNVYSEVLRSLQKA 298

```

QY 312 SKMDLVLRKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKQRQDKKLODQ 370
DB 299 SKMKELVLKKSSELEIYRKTTHMVGSDGAMNIAIEAIESGAVNDADAVALQELRIATV 358
QY 371 KEAFSRKDIILKVERWQACBEEAWLEDYNDKNNRAGRGHLLTKRAEKARTLVNKI 430
DB 359 KEAFSRKDIILKVERWQACBEEAWLEDYNDKNNRAGRGHLLTKRAEKARTLVNKL 418
QY 431 PGWVVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKQRQDKKLODQ 490
DB 419 PAMVEALASKTKAWENERGTQ-FSYDGIPLLSMLEEYTLREKELEKRRKQDKKLOQ 477
QY 491 LKAEQALYKSPSPKPLSTTKAPRHSMDGA--NRRLSLGGATWQPPKTDILHS---K 544
DB 478 LMAEQESLYGSKPSPMKNSAKGPKLSCGAPSNRRSLGGTMMQTKTELPHSTKATP 537
QY 545 SVRAAKKTEE-----IGTLPSSSRGLDIAGLPIKKLSFNASTLTRETTP- 589
DB 538 NTRQAKKSRFHQLDQFNHPTNDGFGALSAGRRGGLIDELPSKKQPIALSNGSEVETAV 597
QY 590 -RKPPFAQI--TPGNSVS-----STPVRPITNNTEDE 618
DB 598 MRKPPSPISSGSKSSKNATNLGDMNRKHESMIKTLNHNHTTPVSTPVKSI--STCBE 655
QY 619 NRTTP-KTFTALNPKTMTVTAPMQAMTSPKANKYSATPVS--LVYDKPEVTLQEDIDYS 675
DB 656 NRTPAKAMPFVPSTPSTVSVPMQ-TTTPGPA---VVTYPNSKLV-----ENTHVEIEYS 707
QY 676 FEERLAIYLAR 687
DB 708 FEESRLEFYLR 719

RESULT 2
AAT40494
ID AAT40494 PRELIMINARY; PRT; 730 AA.
AC AAT40494;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative microtubule-associated protein.
GN FGEC561.15.
OS Solanum demissum (wild potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=50514;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R.C., Liu J., Zaborsky J., Tallon L., Baker B.;
RT "Solanum demissum chromosome 11 BAC FGEC561 genomic sequence.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ronning C.M.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC136471; AAT40494.1; -.
SQ SEQUENCE 730 AA; 82876 MW; E4E80EBD8F60D4E5 CRC64;

Query Match 55.2%; Score 1929.5; DB 2; Length 730;
Best Local Similarity 57.1%; Pred. No. 1.5e-79;
Matches 418; Conservative 91; Mismatches 154; Indels 69; Gaps 15;

QY 12 MTTTCDLLELNLVWDEVEGPDTRDRMLLEOECELVYRKYDQANRRAQRLKRIA 71
DB 1 MTTTGLLLHQLIIVDEVEGCDTERDMLFIERECLELVYRKYDQANRRAQRLKRIA 60
QY 72 EGAEALAGSANGEPFVHVRSQNKHLHGLREELNAIVPYLEEMKKKKVWRWNQFVHVIE 131

```

```

DB 61 DSEALAAICAMGERPVHIIKQSKSQGLKAEALRAVLPELEEMKRKSRDKNQFIEVMK 120
QY 132 QIKKISSIRPADFPVFPKVDQSDLSLRKLDLTKDLESLOKESKDRKQVIBHLNSLH 191
DB 121 QITKIKNEI--YRFTSASLVVDESLSLRKLEELHTLHTLQKEKSERLKQVNLHGLTN 178
QY 192 SLCEVIGDIFKQTVYVHPSLDEAGSKULSNNTTIERLAAAANRLREMKIQRMQKLODPA 251
DB 179 SLCSVLGMDFKHTINEVDPNLGESEAKNICDDTIONLAATIQRLQEVKLRMORLODPT 238
QY 252 SSMLLELNLMDTPLEEQOMFONITCNIAASQEIETEPNTLSTDFLNYVESEVLRLEQLKA 311
DB 239 TSMLLELNLMDTPLEEQOMFONITCNIAASQEIETEPNTLSTDFLNYVESEVLRLEQLKA 298
QY 312 SKMDLVLRKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKQRQDKKLODQ 490
DB 299 SKMKELVLKKSSELEIYRKTTHMVGSDGAMNIAIEAIESGAVNDADAVALQELRIATV 358
QY 371 KEAFSRKDIILKVERWQACBEEAWLEDYNDKNNRAGRGHLLTKRAEKARTLVNKI 430
DB 359 KEAFSRKDIILKVERWQACBEEAWLEDYNDKNNRAGRGHLLTKRAEKARTLVNKL 418
QY 431 PGWVVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKQRQDKKLODQ 490
DB 419 PAMVEALASKTKAWENERGTQ-FSYDGIPLLSMLEEYTLREKELEKRRKQDKKLOQ 477
QY 491 LKAEQALYKSPSPKPLSTTKAPRHSMDGA--NRRLSLGGATWQPPKTDILHS---K 544
DB 478 LMAEQESLYGSKPSPMKNSAKGPKLSCGAPSNRRSLGGTMMQTKTELPHSTKATP 537
QY 545 SVRAAKKTEE-----IGTLPSSSRGLDIAGLPIKKLSFNASTLTRETTP- 589
DB 538 NTRQAKKSRFHQLDQFNHPTNDGFGALSAGRRGGLIDELPSKKQPIALSNGSEVETAV 597
QY 590 -RKPPFAQI--TPGNSVS-----STPVRPITNNTEDE 618
DB 598 MRKPPSPISSGSKSSKNATNLGDMNRKHESMIKTLNHNHTTPVSTPVKSI--STCBE 655
QY 619 NRTTP-KTFTALNPKTMTVTAPMQAMTSPKANKYSATPVS--LVYDKPEVTLQEDIDYS 675
DB 656 NRTPAKAMPFVPSTPSTVSVPMQ-TTTPGPA---VVTYPNSKLV-----ENTHVEIEYS 707
QY 676 FEERLAIYLAR 687
DB 708 FEESRLEFYLR 719

RESULT 3
Q9FHM4
ID Q9FHM4 PRELIMINARY; PRT; 707 AA.
AC Q9FHM4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Emb|CAB82688.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000)
DR EMBL; AB018109; BAB08676.1; -.
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.

```

SEQUENCE 707 AA; 80301 MW; 30EAF1E5BCC6B9A CRC64;  
Query Match 54.4%; Score 1900.5; DB 2; Length 707;  
Best Local Similarity 55.1%; Pred. No. 3e-76;  
Matches 397; Conservative 120; Mismatches 148; Indels 55; Gaps 12;  
QY 1 MSSAVKQDLQHMSTTCDLSLLLELVNWDVSGPDPDTRDRMLLEBOECLELVYRRKVDQAN 60  
DB 1 MASVQKDPILQVETTCGSLFELQIWDVSGTETDRDQMLLEBOECLELVYRRKVDQAN 60  
QY 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKQLHGLREELNAIVPYLEEMKKVKV 120  
DB 61 RCRAQLRQAIAEAGEAELAGICSAMGERPVHVRQSDVSGSLKQELGRILLPELEEMKKVKV 120  
QY 121 ERWNOFVHVIEQIKISSIR-PADVPFVKVPVQSDLSLRKLDLTKLESLOKESDR 179  
DB 121 ERNQPFIIVMEQIDSTNDIKQGGELVHSEPIDETNLSMRKLEELHCOQLQVLOKEKIDR 180  
QY 180 LKQVIEHLSLSCVLGIDFKQTVYVHPSLDEAGSKNLSNTTIERLAAANRLREM 239  
DB 181 VETIRKHLCTLSHCSVLGMDNEVVGVNPTLSDPGRSLSDHTIEKLGAAVQKLMV 240  
QY 240 KIORMKQDLQFASMLNLMNDTPEEQOMFONITCNIAAEOEITEPNTLSTDFLNVY 299  
DB 241 KIORMKQDLQFASMLNLMNDTPEEQOYQVHITCNIAAASHEITEANSLEDPIKYV 300  
QY 300 ESEVLRLQOLKASMKDLVLKKAELERHRRRAHLVGEYGAEEFSEIAEAGIDPSLV 359  
DB 301 EAEVRLDEYKASKMELVLKCRSELEECRKTHLLPVSDSAIDQTVIAESGIVDATWV 360  
QY 360 LEQIEAHIAIVTKEEAFSRKDIIEKVERWONACBEEAWLEDYKNDNRYNAGRGHLLTKR 419  
DB 361 LEHLEQHSIKIEEALSRKEILRVEKNLSACDEESWLEYNDRDNRYNAGRGHLLTKR 420  
QY 420 AEKARTLVNKPQWVDVLRKIAAANKRGEKEDFTYDGVSLSSMLDEYMFVROEKEKK 479  
DB 421 AEKARNLVTKLPGWVEALASKTVWEQENGIE-FLYDGIILLSMLEYNILROEREEHR 479  
QY 480 RORDQKQDLQKAEQALYGSFSPSKPLSTKKAAPRSHMGGN-NRRLSISGGATMPPKT 538  
DB 480 RORDQKQDLQKAEQALYGSFSPSKPLGGKKAPRSTGGASNRRLSISGAAMHQTPKP 539  
QY 539 DILHSKSVRAAKTEETIGTLPSSSRGLDIAGPIKKLSFNASTLRETETPRKPPA- --- 594  
DB 540 N-----KKADRHNDGALS-NRRGLDIALPSRKQSNVPMQLSPVRKPFSPIST 591  
QY 595 -----QITPGNSVS-----STPVR---PITNNTDE-----NRTPKTFTA 627  
DB 592 TVVASKANIATTTTQQLPKONAVNEISSPATPIKNNILRNLEBKMTMMQTPKNVAA 651  
QY 628 L--NPKTPMTVTPMQAMTPSLANKVSATPVSILVYDKPVTLOEIDISFEERRLAIVL 685  
DB 652 MIPISTPATVSPVMTAPTFTNN-----ARLMSEKPEV-----VEYSFEERRLAIVL 700

## RESULT 4

Q8LQES PRELIMINARY; PRT; 645 AA.  
AC Q8LQES;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cytokinesis regulating protein-like.  
GN Name=P0478H03.25;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,

RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL; AP003452; BAB92709.1; -;  
DR Gramene; O8LQES3; -;  
DR InterPro; IPR007145; MAP65\_ASE1.  
DR InterPro; IPR009061; Putativ\_DNA\_bind.  
DR Pfam; PF03999; MAP65\_ASE1; 1.  
SQ SEQUENCE 645 AA; 70887 MW; 2DE3181F1B8P8ED8 CRC64;  
Query Match 46.5%; Score 1625.5; DB 2; Length 645;  
Best Local Similarity 61.1%; Pred. No. 7.9e-66;  
Matches 361; Conservative 16; Mismatches 61; Indels 153; Gaps 9;  
QY 1 MSSAVKQDLQHMSTTCDLSLLLELVNWDVSGPDPDTRDRMLLEBOECLELVYRRKVDQAN 60  
DB 1 MSSAVKQDLQHMSTTCDLSLLLELVNWDVSGPDPDTRDRMLLEBOECLELVYRRKVDQAN 60  
QY 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKQLHGLREELNAIVPYLEEMKKVKV 120  
DB 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKQLHGLREELNAIVPYLEEMKKVKV 120  
QY 121 ERWNOFVHVIEQIKISSIRPADVPFVKVPVQSDLSLRKLDLTKLESLOKESDR 180  
DB 121 ERWNOFVHVIEQIKISSIRPADVPFVKVPVQSDLSLRKLDLTKLESLOKESDR 180  
QY 181 KQVIEHLSLSCVLGIDFKQTVYVHPSLDEAGSKNLSNTTIERLAAANRLREM 240  
DB 181 KQVIEHLSLSCVLGIDFKQTVYVHPSLDEAGSKNLSNTTIERLAAANRLREM 240  
QY 241 IORMKQDLQFASMLNLMNDTPEEQOMFONITCNIAAEOEITEPNTLSTDFLNVY 300  
DB 241 IORMKQDLQFASMLNLMNDTPEEQOMFONITCNIAAEOEITEPNTLSTDFLNVY 300  
QY 301 SEVLRLQOLKASMKDLVLKKAELERHRRRAHLVGEYGAEEFSEIAEAGIDPSLV 360  
DB 299 ----- 298  
QY 361 EQIEAHIAIVTKEEAFSRKDIIEKVERWONACBEEAWLEDYKNDNRYNAGRGHLLTKRA 420  
DB 299 ----- 316  
QY 421 EKARTLVNKPQWVDVLRKIAAANKRGEKEDFTYDGVSLSSMLDEYMFVROEKEKK 480  
DB 317 EKARTLVNKPQWVDVLRKIAAANKRGEKEDFTYDGVSLSSMLDEYMFVROEKEKK 480  
QY 481 Q--RDQKQDLQKAEQALYGSFSPSKP-----LSTKKAPRSHMGG- ---A 522  
DB 358 HATPRAEGAVNVKTDARPRATTGVWPAGQPHDARTLARTHNAQAARAGRGACVWDSP 417  
QY 523 NRRLSISGGATMPPKTDLHSKSVRAAKTEETIGTLPSSSRGLDIAGL 571  
DB 418 NTR---DGRTNQVPP-----RAVRPVEPIESIDSLPSVAGELYLVGM 458  
RESULT 5  
Q8LZY0  
ID Q8LZY0 PRELIMINARY; PRT; 648 AA.  
AC Q8LZY0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)





```
Db 367 ISKAEEAFSRKIDLRVKKMSACBESWLEDYNRDQNRYSASRGAGHNLKRAEKARIL 426
QY 427 VNKIPGMVDVLRKIAAKNKGKEDFTYDGVSLSSMLDEYMFVROEKEKKRORDOKK 486
Db 427 VSKIPAMVDVLRKIAAKNKGKEDFTYDGVSLSSMLDEYMFVROEKEKKRORDOKK 485
QY 487 LDQLKAEQEAALYSGKSPSKPLSTKK--APRHSMMGA---NRRLSL 528
Db 486 VQEQPHVEQESAFSTSPSPVSAKTVGPRANNGGANGTHNRRLSL 533

RESULT 7
Q9FEV9
ID Q9FEV9 PRELIMINARY; PRT; 580 AA.
AC Q9FEV9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-JUN-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1a.
GN Name=map65-1a;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ289862; CAC17794.1; -.
DR InterPro: IPR007145; MAP65_ASE1.
DR Pfam: PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 580 AA; 65259 MW; 7519865D362F508A CRC64;

Query Match 37.5%; Score 1311; DB 2; Length 580;
Best Local Similarity 48.5%; Pred. No. 1-2e-51;
Matches 280; Conservative 103; Mismatches 158; Indels 36; Gaps 9;

QY 14 TTCDSLLLELVNVIWDEVEGPDTRDRMLLELEQECLEVVYRKVDQANRSRAQLKRAIAEG 73
Db 14 TTCGSLLOLQIWDVEGETDDBDRNMLLQIQECLEVDVYRKVDQAVKSRHLLQALADA 73

QY 74 EALAGICSANGPVPVHVRQSNQKLHGLREELNAIVPYLEEMKKKKVERWQVHVIEQI 133
Db 74 KVLSRLLSALGE-KTVVGIPKTSGTIKELQALAIAPALEKLEWQKDDRIKEFFDVQSQI 132

QY 134 KTISSERPADFPVKVPVQSDLSLRKLDLTKDLESQKESDRLLKQVIEHLNLSLHSL 193
Db 133 QKISSEIAGTREQVESLTVDSELSLKLDEFOAQIQELQKESERLQKVLVSVTHDL 192

QY 194 CEVLGIDFKQTVYVHPSLDEAG--SKNLSNTTIERLAAAANRLREMKIQRMKQLQDFA 251
Db 193 CAVLGMDFFSTVEVHPSLNDSTGVQSKSISNDTSLSLAKTVLVLKEDKQRLHKLQELA 252

QY 252 SSMLLELNLMDTPLEEQMFQNTICNIAASEQIETPTNTLSTDFLNVSEVLRBOLKA 311
Db 253 TQLIDLNLMDTPEERSESLDFHVTNCNISASVDEVAIPGALALDLIOEAEEVEVRLDQLA 312

QY 312 SKMKDLVLKKALEEHRRRAHLVGEGBYAEFSIEAIEAGADIPSLVLRQIEAHATATVK 371
Db 313 SKMKEIAFKQAELEDIYARAHVEIDTEAAREKIMALIDSGNVDPAELLADMNQIVNAK 372

QY 372 EAFSRKDIILEKVERWQACBEEAWLEDYNDKDNRYNAGRAHLLTKRAEKARTLVNKKIP 431
Db 373 EEAHSRKEILDKEKMAACBESWLEDYNRDQNRYSASRGAGHNLKRAEKARILVNNKIP 432

QY 432 GMVDVLRKIAAKNKGKEDFTYDGVSLSSMLDEYMFVROEKEKKRORDOKKLDQL 491
Db 433 ALVDLSLVAKTRAWEQER-DTFTYDGVPLLAMLDYEMMLRHDREEEKRRLDQKFFEQI 491

QY 492 KAEQEAALYSGKSPSKPLSTKK--APRHSMMGA---NRRLSLGGATMQPPTDKILHKS 545
```

```
Db 492 -SKEETVFGSTSPARPLGPKVTGPR-ANGSANGPTSRRLSLNS-----HQNG 538
QY 546 VRAAKKTBEIGTSLSPSSSRGLDIAGLPIKLSFNAST 582
Db 539 SRSTWKGKRD-----RPIAPLNVYVMT 562

RESULT 8
Q84VU1
ID Q84VU1 PRELIMINARY; PRT; 576 AA.
AC Q84VU1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 65kD microtubule associated protein.
GN Name=map65-1;
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12527379;
RA Chan J., Mao G., Smertenko A., Hussey P.J., Naldrett M., Bottrill A.,
RA Lloyd C.W.;
RT Identification of a MAP65 isoform involved in directional expansion
RT of plant cells.;
RL FEBS Lett. 534:161-163 (2003).
DR EMBL: AJ520103; CAD58680.1; -.
DR InterPro: IPR007145; MAP65_ASE1.
DR Pfam: PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 576 AA; 64771 MW; 1E90E3B570CC6128 CRC64;

Query Match 37.1%; Score 1296; DB 2; Length 576;
Best Local Similarity 47.5%; Pred. No. 5.8e-51;
Matches 275; Conservative 102; Mismatches 156; Indels 46; Gaps 10;

QY 14 TTCDSLLLELVNVIWDEVEGPDTRDRMLLELEQECLEVVYRKVDQANRSRAQLKRAIAEG 73
Db 14 TTCGSLLOLQIWDVEGESDAERDKMLLQLEQECLEVDVYRKVDQAVKSRHLLQSLADG 73

QY 74 EALAGICSANGPVPVHVRQSNQKLHGLREELNAIVPYLEEMKKKKVERWQVHVIEQ- 132
Db 74 QVELHSLMSALGE-KTVVGIPNTSGTIKEQLSALAPTLEQLWKQDKERKFSDFVQSQI 132

QY 133 QNDMCQITGTIEKVGSPV-----VDESDLSLKKLDBFOAKLQELQKESDRLLK 181
Db 183 VIEHLNLSLHSLCEVLGIDFKQTVYVHPSLDEAG--SKNLSNTTIERLAAAANRLREMK 240

QY 182 VLEFVSTVHDLCAVLRMDFFSTVEVHPSLDDSIGVQSKSISNDTSLSLAKTVLALDEBK 241
Db 241 IORMOKLQDFASSMLLELNLMDTPLEEQMFQNTICNIAASEQIETPTNTLSTDFLNVYE 300

QY 242 EQLRLKQLBLAKQLDLNLMDTSEERSLDFHVTNCNVSASVDEVITPGALALDLIEQAE 301
Db 301 SEVLRLEQLKASKMKDLVLKKALEEHRRRAHLVGEGBYAEFSIEAIEAGADIPSLVL 360

QY 302 VEVERLDQLKASRMKEIAFKQAELEEEIYVHAHVEIDSEAAERIMALIDSGDVEPAELL 361
Db 361 EQIEAHIAIVKEASRDKIILEKVERWQACBEEAWLEDYNDKDNRYNAGRAHLLTKRA 420

QY 362 DDMDNQIVVKAKEALSRLDKVEKMSACBESWLEDYNDKDNRYNAGRAHLLTKRA 421
Db 421 EKARTLVNKKIPGMVDVLRKIAAKNKGKEDFTYDGVSLSSMLDEYMFVROEKEKKR 480

QY 422 EKARILVNNKIPALVLSLVAKTRAWEERG-VSFTYDGVPLLAMLDYEMMLRHDREEEKR 480
Db 481 ORDQKKLQDLK-ABQEAALYSGKSPSKPLSTKKAPRHSMMGA-----NRRLSLGGATMQP 535
```

```
Db 481 MKDQKHFHEQTAAEQETAFTARPSPARPVGIKKAAGSVRANGANGTTPNERLSINA----- 535
QY 536 PKTDILH---SKSVRAAKTEBEIGTLP-SSSRGLDIAG 570
Db 536 -----HONGSRSRINKDGRRESMRAAAPLKEDAGSCISG 568

RESULT 9
Q9FEV8 PRELIMINARY; PRT; 584 AA.
AC Q9FEV8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1b.
GN Names=map65-1b;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ289863; CAC17795.1; -
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 584 AA; 65886 MW; 9507821B1A3D6835 CRC64;
```

```
Query Match 36.6%; Score 1279.5; DB 2; Length 584;
Best Local Similarity 48.9%; Pred. No. 3.3e-50;
Matches 267; Conservative 101; Mismatches 161; Indels 17; Gaps 7;

QY 14 TTCDLSLLELVINWDEVGEPDTRDRMLLEBQCELVYRRKVDQANRSRAQLRAIAEG 73
Db 16 TTGCSLLQQLQIWDVEVGQTDDEWHEMLLIQIDRECLDVYKRVQDAVKRAHLLEALADA 75

QY 74 EAEIAGICSGAMGEPVHVRSQNKHLGRLBELNAIVPYLEMKKKVWRNQFVHVIEQI 133
Db 76 KIELCRLLAALGE-KTYAGIPEKASGSIKEQLAALAPALENLWKQEDRVKEFFNVQAI 134

QY 134 KXISSEIRPADVPFVKPVQDLSLRKLDLTDLKLESQKESDRKQVIEHLSLHSL 193
Db 135 EKISSEIAGISEHVNPKNVDESLSLKLDFQALQELQEKSERLHKVLEFVSTIIDL 194

QY 194 CEVLGIDFKQTVYVHPHSLDEAEG--SKNLSNTTIERLAAAANRLREMKIORMQKLDQFA 251
Db 195 CAVLGLDFPFTITEVHPHSLNDSTGVQSOSISNDTSLNSLRTVLALMEDKKQRLQKQLA 254

QY 252 SSMLELWNLMDTPLEEQOMFONITCNIAAEEQTEPTNTLSTDFLNYVESEVLRLEQLKA 311
Db 255 TQLIDLWNLMDTPEERSLFDHVCCNISASVDEVSIPGALALDLTQEAQVEVERLDQLKA 314

QY 312 SKMKDLVLKKAELREHRRRAHLVGEEGYAEFEFSIEAIEAGALDPSLVLEQIEAHIAIVK 371
Db 315 SRMKIEAFKQAVLEIEFAHAIIEIDSEAAQKIMLIESGNIETPAELLTDMQIVKAN 374

QY 372 EAFSRKDILEKVERQWNAACEEAWLEDYNDKDDNRYNAGRAHLLTKAEKARTLVNKKIP 431
Db 375 EEAHSRKEILEKVEKMAACEEESWLEDYNDQNRYNASRGHNLNKAERKARILVNKKIP 434

QY 432 GMVDVLRITKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRRDQKLDQOL 491
Db 435 ALVDSLIAKTRAEQDR-DTTFYDGVFLAMLDEYMWLRHDEERRLRDQKFFHEQI 493

QY 492 KAEQALYGSKPSKPLSTKK--APR---HSMGGANRLSLGG-----ATMOPPKTDI 540
Db 494 SKEPTTFGSPSPARPLGTQKVCPRANGSANGNRRLSLNSHONGSRSTTKDGRD- 552
```

```
QY 541 LHSKSV 546
Db 553 -HSRKV 557

RESULT 10
Q9FEV7 PRELIMINARY; PRT; 582 AA.
AC Q9FEV7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1c.
GN Name=map65-1c;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ289864; CAC17796.1; -
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 582 AA; 65474 MW; 28C05FDEAAB9A975 CRC64;
```

```
Query Match 36.3%; Score 1268.5; DB 2; Length 582;
Best Local Similarity 48.7%; Pred. No. 1e-49;
Matches 266; Conservative 100; Mismatches 163; Indels 17; Gaps 7;

QY 14 TTCDLSLLELVINWDEVGEPDTRDRMLLEBQCELVYRRKVDQANRSRAQLRAIAEG 73
Db 14 TTGCSLLQQLQIWDVEVGQTDDEWHEMLLIQIDRECLDVYKRVQDAVKRAHLLEALADA 73

QY 74 EAEIAGICSGAMGEPVHVRSQNKHLGRLBELNAIVPYLEMKKKVWRNQFVHVIEQI 133
Db 74 KIELCRLLAALGE-KTYAGIPEKASGSIKEQLAALAPALENLWKQEDRVKEFFNVQAI 132

QY 134 KXISSEIRPADVPFVKPVQDLSLRKLDLTDLKLESQKESDRKQVIEHLSLHSL 193
Db 133 EKISSEIAGISEHVNPKNVDESLSLKLDFQALQELQEKSERLHKVLEFVSTIIDL 192

QY 194 CEVLGIDFKQTVYVHPHSLDEAEG--SKNLSNTTIERLAAAANRLREMKIORMQKLDQFA 251
Db 193 CAVLGLDFPFTITEVHPHSLNDSTGVQSOSISDDTSLNSLRTVLALTEDKKQRLQKQLA 252

QY 252 SSMLELWNLMDTPLEEQOMFONITCNIAAEEQTEPTNTLSTDFLNYVESEVLRLEQLKA 311
Db 253 TQLIDLWNLMDTPEERSLFDHVCCNISASVDEVSIPGALALDLTQEAQVEVERLDQLKA 312

QY 312 SKMKDLVLKKAELREHRRRAHLVGEEGYAEFEFSIEAIEAGALDPSLVLEQIEAHIAIVK 371
Db 313 SRMKIEAFKQAVLEIEFAHAIIEIDSEAAQKIMLIESGNIETPAELLTDMQIVKAN 372

QY 372 EAFSRKDILEKVERQWNAACEEAWLEDYNDKDDNRYNAGRAHLLTKAEKARTLVNKKIP 431
Db 373 EEAHSRKEILEKVEKMAACEEESWLEDYNDQNRYNASRGHNLNKAERKARILVNKKIP 432

QY 432 GMVDVLRITKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRRDQKLDQOL 491
Db 433 ALVDSLIAKTRAEQDR-DTTFYDGVFLAMLDEYMWLRHDEERRLRDQKFFHEQI 491

QY 492 KAEQALYGSKPSKPLSTKK--APR---HSMGGANRLSLGG-----ATMOPPKTDI 540
Db 492 SKEPTTFGSPSPARPLGTQKVCPRANGSANGNRRLSLNSHONGSRSTTKDGRD- 550

QY 541 LHSKSV 546
Db 551 -HSRKV 555
```

Qy	547	RAAKTTEETGL-SPSS--SRGLDIAGLPIKK	575
Dd	539	GSKSRDLSLNKIASPSNIVANTKDDAASPVS	570
RESULT 12			
Q9SIS3		PRELIMINARY;	PRT; 608 AA.
ID	Q9SIS3		
AC	Q9SIS3;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Hypothetical protein At2G01910.		
GN	NameAt2G01910;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OX	eurosid II; Brassicales; Brassicaceae; Arabidopsia.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RA	Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,		
RA	Barnstead M.E., Meason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,		
RA	Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,		
RA	Fraser C.W., Venter J.C.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RA	Town C.D., Kaul S.;		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.		
EMBL:	AC007069; AAD21782.1; -.		
DR	PIR; F84430; F84430.		
DR	InterPro; IPR007145; MAP65_ASE1.		
DR	Pfam; PF03999; MAP65_ASE1.1.		
KW	Hypothetical protein.		
SQ	SEQUENCE 608 AA; 5AE1314D0D311413 CRC64;		
Query Match	34.58;	Score 1208;	DB 2; Length 608;
Best Local Similarity	39.98;	Pred. No. 6e-47;	
Matches 256;	Conservative 131;	Mismatches 185;	Indels 70; Gaps 9;
Qy	13	STTCDSLLLEINVIWDVGEPTTRDRMLLELEQECLLEVYRRKYVDQANRGARLKIAE	72
Dd	15	NITCNLLRELQIKWIWEIGETEKDRMLELERECQLIQRYKVDEAANSKAKHLQSVAS	74
Qy	73	GEAELAGICSGAMG----EPVHVVRQSNQKLHGLEELNAIVPYLEEMKKVKVEENQVFH	128
Dd	75	IEAEVASLMALGVNLINSPILDKGSK---SLKEKLAAVTPVLEEURLIOKEERMKOFSD	131
Qy	129	VIEQIKKISSEIRPAD-----FVPFKVPVDQSLSLRKLDELTKDLSELOKEKSDRLKQV	183
Dd	132	IKAQEKISEIGSYSHLNKANMISITLEEQDLTLRNINEYQTHLTLOKEKSDRLNKV	191
Qy	184	IEHLNSHSLECVLGIDFKQTVEVHPSL--DEAGSKNLSNTTIERRAAANRLREMKI	241
Dd	192	LGYNVHEVALCGVLGVDFQTSVAVHPSLRHRTDQEQSTNISDSTLEGLEHMIIQKLKTERK	251
Qy	242	QRMKQLDPASSMLELNWLMDTPLEEQOMQNICTNAASEQETEPNTLTSTDFLNFVYES	301
Dd	252	SRFOKLKDVMASLPFLNWLMDTPQEDRTKFGKVTYVVVRSSANITEPGILTSTETIEQVST	311
Qy	302	EVLRLSQLKSMKMDLVLKKAEEEHRRRAHLVGBEGYAEFEFSIEAIEAGAIDPVLVE	361
Dd	312	EVDLSKLKSARKMELVNKRSELEDLCRLTHIQPDITSFAEKSTALIDSLVDPSSELLA	371
Qy	362	QIEAHIAIVKEEAFSRKDIILEKVERWNQACEEEEEAWLDYNKDDNRYNAGRAHULTKAAE	421
Dd	372	NIEWQINKIKDEAQRSKQINDRIDRWLSACEEENWLBSEYNLDENRYISAGROGHVNLKAAE	431
Qy	422	KARTLVNKKIPGMVDVLTATKAANKVNGEKEDFTYDGVSLSSMLDYSVMFVROKEQEKKRQ	481
Dd	432	RARYTNKIPGMVDVLTATKTLVNEEDMQK-SFLYDGVRLVNILEDNYKULTRQKEEEKURY	490

RESULT 11	Q8L5G3	PRELIMINARY;	PRT;	578 AA.
ID	Q8L5G3			
AC	Q8L5G3			
DT	01-OCT-2002 (TrEMBLrel. 22, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Microtubule-associated protein MAP65-1a.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI TaxID=3702;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22089475; PubMed=12093376;			
RA	Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,			
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;			
RT	"Full-length messenger RNA sequences greatly improve genome			
RT	annotation.";			
RL	Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,			
RA	Feldmann K.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY085430; AAM62657.1; -			
DR	InterPro; IPRO07145; MAP65_ASE1.			
DR	Pfam; PF03999; MAP65_ASE1; 1.			
SQ	SEQUENCE 578 AA; 65204 MW; A8544528F8C2DE CRC64;			
Query Match 35.4%; Score 1236.5; DB 2; Length 578;				
Best Local Similarity 44.9%; Pred. No. 2.9e-48;				
Matches 257; Conservative 122; Mismatches 166; Indels 27; Gaps 8;				
Qy	15	TCDSLLLELVINWDEVGPDTRDRMLLELEBQECLEVVYRRKVDQANRSRAQLRAIAEGE	74	
Db	15	TCGTLQLQEIWDVSGSDSEERDKLLIQIBEECLNVYKKVLEAAKSRAELLQTLSDAT	74	
Qy	75	RELAGICSAEGEPPVHVHQSNQKHLGLEELNATVPYLEEMKKKKVERWNQFVHVIOIK	134	
Db	75	VELGNLTALGKES-YIDIPKTSGTGKEQSLAIPALEQWQKEEVRAFSVDVQSQIQ	133	
Qy	135	KISSEIR-----PADFVPFVKPVQDQLSKRLDELTKLLESLQEKESDRLLKQVIEHLN	188	
Db	134	KICEIAGGLNGPH-----VVDETDLSLKRLLDFQKQLQEKESDRLLQKLVFVS	186	
Qy	189	SLHSICEVLGIDFKQTVVEVHPSLDEARG--SKLSNTTIERLAAAANRLREMKIQMQK	246	
Db	187	TVHDLCAVLRDLFTSTVTEVHPSLDEANGVQTKSISNETLARLAKTVTLKEDKMQRLKK	246	
Qy	247	LQDFASSMLELWNLMDTPTLEQQMFQNTICNIAASEQEIITPNTLTSTDFLNVESEVLR	306	
Db	247	LQELATQTLDNLNMDTSDEREFLDHTVISNISASVHEVTASGALALDLIEQAQEVVDRL	306	
Qy	307	EQLKASKMKDLVLKKALEEHRRRAHLVGEEGVAAEFISIAIEAGADDPDLVLEQIEAH	366	
Db	307	DOLKSSRMKEIAFKQSELEIYARAHIEIKPEVVRRIRMSLIDAGNTEPTELLADMDSQ	366	
Qy	367	IATVKEEAFSPKDIILEKVERWQACEEAWLEDYKDDNRNRYNAGRGALHTLKRAEKARTL	426	
Db	367	JAKAKEEAFSRKEIILDRVEKWMKSAACEESWLEDYNRDQNRYSASRGALHNLKRAEKARIL	426	
Qy	427	VNKIPGWDVDRTKTAAMKNRGEKDFDYDGVGSLSSMLDEYMFVHQEKEQEKORDQKK	486	
Db	427	VSKITAMVDTLTAKTRAMEEENSM-SFYEDGVPLAMLDDEYTWLRQEEDEDEKRLKEQKK	485	
Qy	487	LQDLKAEQALYLGSKPSFKPLSTKXAPRHSMMGGANRRLSLGGATWQPPKTDILHKSIV	546	
Db	486	QFOEPTHOESAFSGSKSPARVPVSAKKP-----VGTFRVGGGLNETFMRLSNMNSQON	538	



QY 298 VVESEVRLLEQKAGKMDLVKKAELEEHRRRAHLV-GBEGYAEFSIEAIEAGADP 356  
Db 300 EVKALCRLELKNWKMELVLKSESELEICRTHIVLEEDIAENVIRKIESGVNP 359  
QY 357 SLVLEQIEAHIAITVKEEAFSRKDIILEKVERWQNAEEBAAWLEEDYNDKDNRYNAGRGHLLT 416  
Db 360 ENILQIEVRACKVKEEALSRLKEILKADKWLNACEEENWLEEVNQDENRYNAGRGSHLI 419  
QY 417 LKRAEKARTLVNKKIPGMVDVLTTKIAAWKNRGEKEDFTYDGVSLSSMLDEYMFVROEKEQ 476  
Db 420 LKRAEKARLVNKKIPAMVEALASKITIWESEK-EYEFUFDESPTNARVYGTQRRRTIR 478  
QY 477 -BKQRQDKLQDOLKAEQALYGSKPSKPLSTKKAPRHSMG 520  
Db 479 TPQEGSETSRSSDIRQRKCY-----EAPKREKG 510

RESULT 15

Q6Z113 PRELIMINARY; PRT; 589 AA.  
AC Q6Z113;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Putative microtubule-associated protein MAP65-1a.  
GN Name=OJ1008\_D06.15;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP04040; BAD12872.1; -  
DR InterPro; IPR007145; MAP65\_ASE1.  
DR InterPro; IPR009061; Putativ\_DNA\_bind.  
DR Pfam; PF03999; MAP65\_ASE1; 1.  
SQ SEQUENCE 589 AA; 66465 MW; 42AA084B3CD54A3 CRC64;

Query Match 33.8%; Score 1182.5; DB 2; Length 589;  
Best Local Similarity 42.9%; Pred. No. 8.3e-46;  
Matches 255; Conservative 133; Mismatches 176; Indels 31; Gaps 12;

QY 1 MSAAVKDQLHQMSTTCDLSLLELVNVIWDEVEGPDTRDRMLLEBOECLEVYRKKVDQAN 60  
Db 1 MSALLRE-----TSCGSLILKQLQSVNDEVGESEEDRDKVLFDQOECLDVKYKVDQAT 54  
QY 61 RSPAQLKATAEAGEAELAGICSAMGPPVHVRSQNKHLGRLREELNAIVPYLEEMKKKV 120  
Db 55 KSRDLLQALDYKTELARIALSALGSKSDI-SPEKTARTIKQLTAIAPTLEQLGKKKK 113  
QY 121 ERNMQFVHVEIQIKTSSEIRPADFPFKV---PVDQSLSLRKLDLTNLESLOKES 177  
Db 114 ERIKELANTQSRIEQIRGEIAGTLEMCQQVALPQINEDDLTVRKLRREFQLQLELEKES 173  
QY 178 DRLLQVIEHLNLSLCEVLGIDFKQTVVHVPSLDEAGS--KNLSNTTIERLAAAANR 235  
Db 174 RLLEKLVHVGWVHDLNVLGMDFFRTITQVHSSLDSDSIGNHEKNISNETLSKLDRTIGT 233  
QY 236 LREMKTRMQKLODPASSMLELNLMDTPLEEQMFQNTICNTAASEQIEPTNTLSTDF 295  
Db 234 LNEDKRLRLKLEQLATQLYDLMDLMDTPVEERSLFDHVSNCNRTATVEEVVPGALADV 293  
QY 296 LNVVESVRLLEOLKASKMDLVKKAELEEHRRRAHLVGBEGYAEFSIEAIEAGAD 355  
Db 294 IDQAQTEVERLDQKYSKMKIEAFKKQAILLEDIYASTHVLDTAVAHEKIQALIESGME 353  
QY 356 PSLVLEQIEAHIAITVKEEAFSRKDIILEKVERWQNAEEBAAWLEEDYNDKDNRYNAGRGHLL 415  
Db 354 PSELIADMDSQILKAKEALSRLKEILKADKWLNACEEENWLEEVNQDENRYNAGRGSHLI 413

QY 416 TLKRAEKARTLVNKKIPGMVDVLTTKIAAWKNRGEKEDFTYDGVSLSSMLDEYMFVROEKE 475  
Db 414 NLKRAEKARILVSKIPALVETLVAKTRAWEEHHGL-PPMYDGVSLLAAMLDEYVILROBRE 472  
QY 476 QEKQRQDKLQDOLKAEQALYGSKPSKPLSTKK-APRHSMGGA-----NRRLSL 528  
Db 473 EEKRMREQRKQTQLLNIDREGPFGTRVNVRYVTSAKKVAGTKPNGGASNGTFSRRLST 532  
QY 529 GGATWQPPKTDILHLSKV-RAAKKTEEIGTLPSSSRGLDIAGLPKLSFNAST 582  
Db 533 G-----NQLNESKSTGRSAGDKDKGA-SKNTATSLNEAA-PADKEAADST 577

Search completed: November 20, 2004, 00:42:10  
Job time : 109 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:14:11 ; Search time 25 Seconds

(without alignments)  
2655.593 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKDLQHQWTTCDLSL.....DIDSFERRLAIVLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79;\*

1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362.5	39.0	648	T47895	hypothetical prote
2	1208	34.6	608	F84430	hypothetical prote
3	1042	29.8	587	E84808	hypothetical prote
4	1004	28.7	473	T04799	hypothetical prote
5	905	25.9	592	E86404	hypothetical prote
6	259.5	7.4	885	S59650	anaphase spindle e
7	230	6.6	880	F75103	conserved hypothet
8	228.5	6.5	1972	A41604	myosin heavy chain
9	222	6.4	1827	T16270	hypothetical prote
10	221	6.3	1780	T17272	hypothetical prote
11	221	6.3	1938	TJCS421	smooth muscle myos
12	221	6.3	1972	JC5420	smooth muscle myos
13	217	6.2	2116	A26655	myosin heavy chain
14	214	6.1	886	H69378	conserved hypothet
15	213.5	6.1	1963	1 MWK	myosin heavy chain
16	211	6.0	1927	A59236	embryonic muscle m
17	210.5	6.0	830	T00029	Miranda protein -
18	210.5	6.0	2020	T21174	hypothetical prote
19	209	6.0	848	A44972	paramyosin - nemat
20	209	6.0	1937	T138055	myosin heavy chain
21	208	5.9	936	T39083	myosin heavy chain
22	208	5.9	1256	T26101	hypothetical prote
23	208	5.9	1690	T13030	microtubule bindin
24	207.5	5.9	876	A23767	myosin heavy chain
25	207	5.9	676	S00084	myosin heavy chain
26	207	5.9	741	T39082	myosin heavy chain
27	207	5.9	2101	A42184	nuclear mitotic ap
28	206.5	5.9	1175	C35815	myosin heavy chain
29	206.5	5.9	1201	A35815	myosin heavy chain

ALIGNMENTS

RESULT 1

T47895 hypothetical protein T4C21.250 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T47895

R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; S.

.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479

A:Accession: T47895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-648 <CHO>

A:Cross-references: UNIPROT:Q9LZV0; EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21

C:Genetics:

A:Map position: 3

A:Introns: 61/3; 145/3; 214/3; 266/3; 320/1; 370/3; 400/1; 424/3; 449/3; 535/1

A:Note: T4C21.250

Query Match 39.0%; Score 1362.5; DB 2; Length 648;

Best Local Similarity 45.1%; Pred. No. 3.6e-57;

Matches 307; Conservative 124; Mismatches 184; Indels 65; Gaps 18;

QY 30 VGEPTTRDRMLLELEOECELVYRKVDQANRSRAQLRKAIAEGEAEAGICSAMGEPV 89

Db 1 MGETEDEKADSLADIEKECLSVYKRKVEASRGKANLLKEIAVGRABIAAIGSMGQEI 60

QY 90 HVRQSNOKL-HGLREELNAIVPVLEEMKKKQVWRNQFVHVIEQIKKISSEI-PPADFPV 147

Db 61 H--SNSRIGENLKELENNVQLDGLRKRKAERMIRFNEVIDQLKLSLQLCNPTDYLK 117

QY 148 FKVPVDQSLSRKLDLTKLESQKESDRLKQVIEHLSLSLCEVLGIDFQKTVYE 207

Db 118 -KFAAEETDLSLQRLLEELRSQGLQNEKSKRLEEVECLLKTLSLCSVLGDFKGMIRG 176

QY 208 VHPSLDEAGSGKNLSNTTIERLAAANRRREMKIORMOKLODFASMLNLWMDTPLEE 267

Db 177 IHSSLVDS-NTRDVSRSSTLDKLDMMITVNRKALQRMQKVQDLAVLLELWNLDTPAEE 235

QY 268 QMFMQITCNIAASEQIEPTNTLTDFLNVYSEVLRLQKASMKDKLVKKKAELEE 327

Db 236 QKIFHNVTLSIALTESEITEANILSVASIKRVEDEVIRLSKITIKIKEVILAKRLEEE 295

QY 328 HRRRAHLVGEEGVAEEFSIEAIEAGAIIDPSLVLEQIEAHIAIVKBEAFPSKQILEKVERW 387

Db 296 ISRKMHMATEVLKSENFSEVATIESGVKPEQLLEQIDSEIAKVKBEASSRKEILEKVEKW 355

QY 388 QNACBEEAMLEEDYKDDNRYNAGRGAAHLTKRAEKARTLVNKIPGVVDVLTAKTAAWNE 447

Db 356 MSACBEEWLEEDYKDDNRYNAGRGAAHLTKRAEKARLLVNLKPLGMVEALTAKVTAWENE 415

NMDA receptor-bind  
tpr protein - huma  
early endosome ant  
myosin heavy chain  
myosin heavy chain  
ATPase involved in  
hypothetical prote  
centromere protein  
P115 homolog - Met  
transport protein  
nonmuscle myosin I  
myosin heavy chain  
paramyosin - nemat  
myosin heavy chain  
myosin heavy chain  
myosin heavy chain





Db 423 IFVFKQFNKKFWALLFSPKQKFNSIIDYCNSTTHSMVDVLTITTKVAKWEKRG-VPF 481  
QY 454 TYDGVLSLMLDEYMFVROEKQEKQRDQKQDOLKAOEALYGSFSPSKPL----- 509  
Db 482 LCDKQPLLOTLEDDIVIRAQREERKQFREQRLQGLQGLATEKEAKYSAXKPLGOSL 541  
QY 510 ---STKAP--RHSMGANRLSLGG 530  
Db 542 NTDNVTKTPIGRIGNTPGRSVTSGG 567

RESULT 4  
T04799  
hypothetical protein F10M23.100 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04799  
R:Bevan, M.; Lecharny, A.; Chedford, F.; Krivitzky, M.; Kreis, M.; Hoheisel, J.; Mewes, H.  
submitted to the Protein Sequence Database, February 1999  
A:Reference number: Z15385  
A:Accession: T04799  
A:Molecule type: DNA  
A:Residues: 1-473 <BEV>  
A:Cross-references: UNIPROT:Q9SZ16; EMBL:AL035440  
A:Experimental source: cultivar Columbia; BAC clone F10M23  
C:Genetics:  
A:Map position: 4  
A:Introns: 24/3; 92/3; 184/3; 256/3; 308/3; 412/3  
A:Note: F10M23.100

Query Match 28.7%; Score 1004; DB 2; Length 473;  
Best Local Similarity 46.9%; Pred. No. 1.9e-40;  
Matches 204; Conservative 94; Mismatches 113; Indels 24; Gaps 4;

QY 15 TCDLLELNLVWDEYGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLKATAEGE 74  
Db 15 TCGTLQKLQEIWDEYGESEDRDKLLQIEEBCLNVYKKVLAAKSRAELLQTLSDAT 74  
QY 75 AELAGICSANGEP-----PVHVRQSNQKLGRLBELNAIVPYLEEMKKKVERWQ 125  
Db 75 VELSNLTALGKSEYIDICDSMSLFLQPDKTSGTIKEQLSAIPALEQLWQOKEERVRA 134  
QY 126 FVHVIEQIKKISSEIR-----PADFVPKVPVDQSLRLKDLTKDLESLOKESKSR 179  
Db 135 FSDVQSQIKICEIAGLNGPH-----VVDETLSLKLDDFKRLQELQKESKR 187  
QY 180 LKQVIEHLNLSLCEVLGIDFKQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLR 237  
Db 188 LQKVLFEVSTVHDLCAVRLDLFLSTVEVHPSLDEANGVQTKSISNETLARLAKTVLT 247  
QY 238 EMKIQRKQKLDQPASSMLELWNLMDTPEBQQFQNTCNIAASEOEITEPTNLSTDFLN 297  
Db 248 EDKQRLKKLQELATQLTDLNLMNLDTSDEERELFDHVTNISAVHEVTASGALALDIE 307  
QY 298 YVESEVLRLEOLKASKMKDLVLKKAELREHRAHLVGEGVAEBSFSAIAGADIPS 357  
Db 308 QAEVEVDRLDQLKSSRNKEATFKQKELEIYARAHIEIKPEVVRERIMSLIDAGNTEPT 367  
QY 358 LVLEQIEAHATVKEAFSRKDIKVERWQNAACEEAWLEDYKDDNRYNAGRGAHLTL 417  
Db 368 ELLADMDSQIAKAKEAFSRKEILDRVKKWMSACEEAWLEDYKDDNRYNAGRGAHLNL 427  
QY 418 KRAEKARTLVNKKIPG 432  
Db 428 KRAEKARILVSKITG 442

RESULT 5  
E86404  
hypothetical protein F13K9.3 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86404  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, K.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; A.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroo, J.S.; Maity, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: E86404  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-592 <STO>  
A:Cross-references: UNIPROT:Q9C7G0; GB:AE005172; NID:gl1024863; PIDN:AAG26947.1; GSPDB:  
C:Genetics:  
A:Map position: 1

Query Match 25.9%; Score 905; DB 2; Length 592;  
Best Local Similarity 38.6%; Pred. No. 1.2e-35;  
Matches 210; Conservative 117; Mismatches 195; Indels 22; Gaps 10;

QY 23 LNVWDEYGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLKATAEGEALAGICS 82  
Db 57 LNMWDEYGEDEKFEREKVLLDIEQCEVAYRRKVDHANVSRHLHQELAESEALTHFL 116  
QY 83 ANGEPPVHVRQSNQKLGRLBELNAIVPYLEEMKKKVERWQFVHVIEQIKKISSEI-- 140  
Db 117 CIGERSVPGR-PEKGGTLREQLDSIAPALREMRUKDERVKQFVRSVGEIKIKISEIAG 175  
QY 141 -RPADFVPKVPVDQSLRLKDLTKDLESLOKESKDRDLKQVIEHLNLSLHSLCEVLGI 199  
Db 176 RSTYEDSTKITDNDLSNKKLEVEQNELHLHDEKNERLQKVDIYICARDLSATLGT 235  
QY 200 DFKQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLREMKIQRKQKLDQPASSMLELW 258  
Db 236 EASMITTKIHPSLNDLYGISKNISDDILKCLNGTIVTVSLSEEEKHLEKLEKHLGRALNSLW 295  
QY 259 NLMDFPLEEQMFQNTCNIAASEOEITEPTNLSTDFLNLYVESEVLRLEOLKASKMKDLV 318  
Db 296 NLMDSYEDRQKFFHVIDLLSPASDVCAPGSITDIIQAAEAEVRLDQKASRIKELF 355  
QY 319 LKKKAELEHRRRAHLVGEGVAEBSFI-EAIEAGAIDPSLVLEQIEAHATVKEAFSR 377  
Db 356 IKKQLEDTCNMHW--ETPTENGNTNLVDSGEVDHVDLLAAMDKEIARAKEEASR 413  
QY 378 KDILEKVERWQNAACEEAWLEDYKDDNRYNAGRGAHLTLKRAEKARTLVNKKIPGVDVL 437  
Db 414 KGIIEKVDRLMLASDEERWLEEDYDQDENRYSVSRNAHRLRAERARITVSKISGLVESI 473  
QY 438 RTKIAAKWNERKEOFTYDGVLSLMDYFMFVROEKQEKQRDQKLO-DQLKAEQ 496  
Db 474 LVKAKSWEVERQKV-FLYNEVPLVAMLQYKNLQKREMEKQKRLREKKMSIDPQVAEGD 532  
QY 497 ALYGKSPSPKPLSTTKAPRHSWGGA-----NRLSLGGATMQPPKPTDILHKSVRAA 549  
Db 533 NYFMARPAS-----SNPRISNRSMNGGFGSGSPINRKYSGGFNTNNNTALGTSIRRES 587  
QY 550 KXTE 553  
Db 588 RKSE 591

RESULT 6  
S59660  
anaphase spindle elongation protein ASE1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein O2806; protein YOR058C  
C:Species: Saccharomyces cerevisiae  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
C:Accession: S59660; S66941  
R:Pellman, D.; Fink, G.R.

submitted to the EMBL Data Library, January 1995  
 A:Description: Yeast microtubule-associated proteins required for anaphase spindle elong  
 A:Reference number: S59660  
 A:Accession: S59660  
 A:Molecule type: DNA  
 A:Residues: 1-885 <PEL>  
 A:CROSS-references: UNIPROT:P50275; EMBL:U20235; NID:g972941; PIDN:AAA75026.1; PID:g9729  
 R:Bohn, C.; Bolotin-Fukuhara, M.; Daigman-Fornier, B.; Dang, D.V.; Valens, M.  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S66929  
 A:Accession: S66941  
 A:Molecule type: DNA  
 A:Residues: 1-885 <BOH>  
 A:CROSS-references: EMBL:274966; NID:g1420196; PIDN:CAA99251.1; PID:e252338; PID:g142019  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:ASE1  
 A:CROSS-references: SGD:S0005584; MIPS:YOR058C  
 A:Map position: 15R

Query Match 7.4%; Score 259.5; DB 2; Length 885;  
 Best Local Similarity 20.7%; Pred. No. 5.8e-05;  
 Matches 173; Conservative 150; Mismatches 288; Indels 225; Gaps 38;

QY 1 MSSAVKQDLQHMTTCDLSLLLELVIVDEVGEPDT---TRDRMLLEQECELVVRRKVD 57  
 DB 86 MDNMYRENFILISKDLEKLENLVYQNIQVSNTEIITKEKIIFTTISNISKQPFQAD 145  
 QY 58 QANRSRAQLKAIAGEAB---LAGICSMGE---PPVHVR-----OSN 95  
 DB 146 E---ELKRLSAENGIEQDILNINLERINDPSGIKTPIDLYIRNAILLQESKTVPOSP 199  
 QY 96 QKLHGLREE---LNAIVPYLEEMKKKVERWQVHVIEQIKKISSIRPAD 144  
 DB 200 KAPLSLLSKAALDAKAFVLSFPLRLDYLSLI---TLKHILQSVKENLPGCLTEAD 255  
 QY 145 -----FVPFKVPVDQS---DLSRLKDELTKDL-----ESL----- 172  
 DB 256 NEAIAEFPPELSTLTAYLLQIENGKGDIGLSMKFIIDNRKIDILKGSFNTINEESVKHNE 315  
 QY 173 -----QKESDLKQVIEHLNLSHLSCEVLGI---DPKQ-----TVY 206  
 DB 316 VIKYIEEYERRFKSVLTAKVSISSICQLGTPLATLIGEDFEQDLRSYGEENSTSEIP 375  
 QY 207 EYHPSLDAEGSKNLSNTTIERLAAANKRREMKIORMKQDQFASSMLELWNLMDTLE 266  
 DB 376 NFHP-VDRERNK--IDITLEKLAHKEADKRLMEQCO-----KLWTRLKISQE 425  
 QY 267 EQOMFONITCNIAASEQITEPTNTLSTDFLNYVESEVLRLEQLKASKMDLV---LKKA 323  
 DB 426 YIKTF-----MRNSSLSTESLGRISKEVMRLKAMKKLIKLLISDWDKIQ 472  
 QY 324 EL-----EHRRAHLVGEYAEESIEAIEAGADPSLVLEQIEAHATVKEEAFS 376  
 DB 473 ELWRTLQVSEERSKFIIVFEELRNSATLQ-----EDLLLETCENELKLEEKULT 525  
 QY 377 RKDILEKVERWONACEEAWLEDYNDKDNRYNAGRAHLLKRAEKARTLVNK-IPGMVD 435  
 DB 526 YKPIKLKISDFESLQEDQEFLEERSKDSRL-LSNSHKILLTEKMKRITRHPFVIN. 584  
 QY 436 VLRTKIAWKNRGKED---FTYDGVSLSSMLDEYVFRQEKQEKQRQDKKIQDQLKA 493  
 DB 585 DLRIKL---EADGLFDQPFLLFKGKPLSEAD-----IQQOEIEAKYPRCVRMRQSKG 636  
 QY 494 EQEALYKSPSPKPLSTKAPRHSWGNGANRLSLGGH--TWQPP-----KTDLH 542  
 DB 637 KCGANKENKVIKNTPKATESIRVDIG-----LNLDANITYKTPSKKTIQGLTKNDLSQ 691  
 QY 543 SKSV-RAAKTTEEIGT-----LSPS-----SSRGLDIAGLPIKLSFNASTLRETETP 589  
 DB 692 ENSLARHMQGTTLKSSPNRRATRLAPTVSRNSKG-----NIERPILNRN--RSSDLS 743  
 QY 590 KRPFAQITPG-NSVSVSTPVRITNTNTEDDEN-----RTPKFTALNPKTPMT-----VTA 638

DB 744 SSPRINHTGHEHVKPRQLPFIPLNKVYDTKGSHTPQLTKKALELLKRESTGTCENVR 803  
 QY 639 PMQAMTSPSLANKVSATPVSIVYDKPEVTL-----QEDIDYSPEE 678  
 DB 804 PERKSSLEDYQAQLSSP-----YKEPEHSIYKLSMSPEKFLQNIQKDISSGDD 854  
 RESULT 7  
 F75103  
 conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay)  
 C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: F75103  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
 A:Reference number: A75001  
 A:Accession: F75103  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-880 <KAW>  
 A:CROSS-references: UNIPROT:Q9U2C8; GB:A0248286; GB:AL096836; NID:g5458366; PIDN:CAB5013  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB0812  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 6.6%; Score 230; DB 2; Length 880;  
 Best Local Similarity 21.1%; Pred. No. 0.0014;  
 Matches 148; Conservative 121; Mismatches 228; Indels 206; Gaps 32;

QY 41 LLEEQECLEVYRKV---DOANRSRAQLKAIAGEAELAGICSANGE---PPVHVRQS 94  
 DB 292 LQEKKE---YKRLKGFDRYESKURRLEKLSKWESELKAEVTKGEKKKERABEI 347  
 QY 95 NOKLGLRLEELNAIVPYLEEMK---KKVERW-----NQFVHVIEQIKKISSEI 140  
 DB 348 REKLSEIEKRLLEELKPYVEELEDKQVQKQIERLKLKGLSPGEVIEKLESLEKTEI 407  
 QY 141 RPADFPVFPVVDQSDLSRLKDLTKDLESLOKESKDRKQVTEHLSHLSCEVLGID 200  
 DB 408 EEA-----IKEITTRIGQMEQEKNERMK-AIEELRKAKGKCPVCGRE 448  
 QY 201 F----KOTVYVHPSLDEAEGSKNLSNTTIERLAAAN-ELREMKIORMKQDQFASSML 255  
 DB 449 LTHEKKELMERY-TLEIKIEELKTTTEERKLRLVNLRLKHLKREFSVMRDIAEQIK 507  
 QY 256 ELNLMMDTPLEEQOMFONITCNIAASEQITEPTNTLSTDFLNYVESEVLRLEQ-LKASK- 313  
 DB 508 EL---ESKLK-----GFNLEELQEKEREFEGLNEEF-NKLKGLLGLERDLKRIKA 554  
 QY 314 -----MKDLVLKKALEEHRRRAHLVCEEGY-----AEPFSIEAIEAGADPS 357  
 DB 555 LEGRRLKIEEKVRKAKBELENHRLQRELGFSEVEELNRIQBLEEPHDFYVKAKESE 614  
 QY 358 LVLEEQIEAHATVKEEAFSRKDLKVKVERWQACEE-EAWLEDYNDKDNRYNAGRAHLL 416  
 DB 615 --LRELKNLEKTELDQAFEMADV---NIEEKEAKLKL---ESKFNE----- 659  
 QY 417 LKRAEKARTLVNKPGMVDVLRTKIAWKNRGKEDFTYDGVSLSSMLDYMVFVRQEKQ 476  
 DB 660 -EYEEKRERLVKLEREVSSLTARLELKK-----SVEQIKATLRKLKEEKEE 706  
 QY 477 EKKRQDOKLQDQKAEQALYGSKPSKPLSTKAPRHSWGNGANRLSLGGATWQPP 536  
 DB 707 REKAEIKKLEKAL----- 721  
 QY 537 KTDILSHKSVRAAKTTEEIGTLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPAQI 596  
 DB 722 -----SKVEDLRKKIKDYKTLA-----KEQALN---RISEIASEIFSEF 757  
 QY 597 TPGNSVSVSTPVRITNTNTEDDENRTPKFTALNPK-TPMTVTA-----PMQAMTTP 646

```
Db 758 TDG-KYSNVIVRA-----EENKT-KLFVVYEGREVPITFLSGGRIALGLAFRLALSM 808
QY 647 SLANKVSVATPSVLVDKPVETVQEDIDYSEFERRLAIFYLAROM 689
Db 809 YLVGRIDL----LILDEPTPFLDEE-----RRRLKLDIMERHL 842

RESULT 8
A1604
myosin heavy chain, smooth muscle, long splice form - rabbit
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A41604; A43501
R:Babji, P.; Kelly, C.; Periasamy, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10676-10680, 1991
A>Title: Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete
A:Reference number: A41604; MUID:92073350; PMID:1961735
A:Accession: A41604
A:Molecule type: mRNA
A:Residues: 1-1972 <BAB>
A:Cross-references: GB:M77812
R:Nagai, R.; Kuro-o, M.; Babji, P.; Periasamy, M.
J. Biol. Chem. 264, 9734-9737, 1989
A>Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cDN
A:Reference number: A33501; MUID:89255535; PMID:2722872
A:Accession: A33501
A:Molecule type: mRNA
A:Residues: 1455-1972 <NAG>
A:Cross-references: GB:J04833; NID:gi65519; PIDN:AAA31407.1; PID:gi65520
A:Experimental source: smooth muscle
A>Note: examination by Southern blotting for the regions of difference between this isofo
active splicing
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F:88-771/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:559-572/Region: actin binding #status predicted
F:633-647/Region: actin binding #status predicted
F:844-1938/Domain: coiled coil #status predicted <COI>
F:844-1284/Region: S2
F:1285-1972/Region: light meromyosin
F:1939-1972/Domain: carboxyl-terminal <CBT>
F:129/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:701,711/Active site: Cys #status predicted

Query Match 6.5%; Score 228.5; DB 1; Length 1972;
Best Local Similarity 19.6%; Pred. No. 0.0048;
Matches 151; Conservative 159; Mismatches 316; Indels 143; Gaps 30;

QY 4 AVKDQLHQ-MSTTCDSLLL-----ELNVIVDEVGPDPTTRDRMLLELEQ---ECLEVVR 53
Db 1150 ALKTELEDLTDTATQQLRAKREQEVTVLKKALDESTRSHEAQVQEMRKHTQVVEELT 1209

QY 54 RYVDQANRSRAQL-----KKAIEGAELAGICSANGPPVHVR----- 92
Db 1210 EQLEQFKRAKANDTKTKQTLKENADLAGSLVGLQAKQVEHKKKLEVLQELQSKCS 1269

QY 93 -----QSNOKLHGLREELNAIVPYLEEMKKKVERWQFVHVIEQIKKISSEIRPADF 145
Db 1270 DGRARALNDKVKHQNEVSTGMLSEAGKAIK-----LAKEVASLGSLQDQTQE 1322

QY 146 VPFKVPVVDQSLRLKLDLTDLKQLESQKESDRL--KQVIE-HLNSHLCEVLGIDFK 202
Db 1323 LLQBEITQKLNVSTKLRLQLEDERNLSQELDDEMEAKONLERHISTLN---IQLSDSK 1377

QY 203 QTVEVHPSLDEAGSKNLSNTTIERL-----AAANRLRMKQRMQKLDQFASML 255
Db 1378 KKLQDFASTVESLEEGKKRFPQKTESLTQOYEEKAAAYDKLEKTKNRLQELDDLVVDLD 1437

QY 256 ELNWLMDTPLEEQMF-----QNITCNIA-----ASEQEITEPTNTLSTDFLNVSEVL 304
```

```
Db 1438 NORQVLSNLEKKQKQKFDQLLAEKKNISSKYADERDRAEAEAREKETKALSARALEEAL 1497
QY 305 RLQQL-KASKMKDLVLKKKAELEHRRRAHLVGBEGYAEFEFSIEAIEAGAI DPSLVLEQI 363
Db 1498 AKBEELERTNKM-----LKAEMEDLVSSKDDVGVKNVHELEKSKRALETQMEEMKQLEBL 1551
QY 364 EAHATVKEEAFSRKDIILEKVERMQNACEEBAWLELDYNDKDDNRYNAGRAH---ITLKRA 420
Db 1552 EDELQAT-EDAKLRLEV--NMQALKVQFERDLQARDEQNEEKKRQLQRLQHEYELEDE 1608
QY 421 EKARTLV-----NKIPGMVDVLRTKIAWKNVERGEDFTYDGVLSLMDLEYMFVQOE--- 473
Db 1609 RKQRALAAAKKLEGG--DLKDLELQADSAIKGREEAIKQLLKLOAQKMFQORELEDARA 1666
QY 474 -----KQOEKKR---ORDOKQLDQLKAEQALYGS---KPSPSKPLSTKKAPRHS 518
Db 1667 SRDEIFATAKENEKAKSLEADLMQLQEDLAAEPARKQADLEKELAEELASSLSGRNA 1726
QY 519 MGGANRRRLSLGGATMQ---PPKTDILHSKSVRAAKKTBIEIGTLPSSSSRGLDIAGLPICK 575
Db 1727 LQDEKKRLEARIQALEEEEBEEQGNMEASDRVRKATQQAELS-----NE 1772
QY 576 LSFNASTLRETETPRKPAQITPGNSVSTPVRPITNNTEDDENRTPKTFTALNPKTPMT 635
Db 1773 LATERSTAQNESARQ---QLERQNKELSKLOEMEGAV---KSKPKSTIAALEAK----- 1822
QY 636 VTAPWQMAWTPSLANKVSVATPSVLVDK--PVTIQEDIDYSEFERRLA 682
Db 1823 -IAQLEEQVEQAREKQAAAKALKORDKKLKEMLLQVE-----DERKMA 1865

RESULT 9
T16270
hypothetical protein F35D11.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16270
R:Fulton, B.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F35D11.
A:Reference number: Z18487
A:Accession: T16270
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1827 <FUL>
A:Cross-references: UNIPROT:Q20042; EMBL:U29391; NID:g868214; PID:g868224; PIDN:AAA6875
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CBSP:F35D11.11
A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1250/3;

Query Match 6.4%; Score 222; DB 2; Length 1827;
Best Local Similarity 18.5%; Pred. No. 0.0088;
Matches 192; Conservative 143; Mismatches 289; Indels 414; Gaps 37;

QY 4 AVKDQLHOMSTTCDSLLLELVINVIDEVBGPDPTRDRMLLELEQ-----QECLEV--YRR 54
Db 492 SUGDVAAMKTDLDKTLRLDML-----ETERDELKKKLETEREQADORDLEIAECRA 544

QY 55 KYVDQANRSRAQLKRAIAEGEA----- 75
Db 545 KLDEMAEKEALREKELAEAFQAITAMEGEGKLNQEQFLESKNKELNLTLDQIESLNSEVEN 604

QY 76 ---ELAGICSANGPPVH---VRQSNOKLHGLREELNAIVPYLEEMKKKVERWQFVHV 129
Db 605 KNEIRNLMTLQEKVEHVIQNVRTSSHQLTATYEEANGSIDILKA-----ELTRL 654

QY 130 IEQIKKISSEIRPADVPFPKVPVVDQSLRLKLDLTDLKQLESQKESDRLKQVIEHLNS 189
Db 655 HEQVNERTKQISEAN-----EKYDDARKNDALLEDVATWQ--EKYEQKMLESEMNR 705

QY 190 LHSCLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNTTIERLAAANRLREM----- 239
```

Db 706 RQEKEREADRALLDLGRFNDKLTNELKQGVTVDSLNEBIEISSLEQNLKSEKERKE 765  
Qy 240 KIQRMOKL-QDPASSMELWNL-MDTPLLEEQMFONITCNIAASEQET-----EPN 289  
Db 766 ELLRMEELQKNEAEKMEYEVKQLQAEKQGVENFQKECEARNELTKIHEMLMEHD 825  
Qy 290 TLSTDFLNVSEVLRL-----OLK 310  
Db 826 QLKVDHL-HTEVEERLEKMKRKELEKLEKNEQDGRDRAWSHNRNPLESSKNEAVTELOER 884  
Qy 311 ASKMKDLVLKXK-----AELEHRRRAHLVGE----- 338  
Db 895 VQKLEVVKEKDEIALRRDLDSHEKSRDLDDKLKRWELTDEKEEDRKKEQKTLNEE 944  
Qy 339 -----GYAF-ESIEAIEAGADP 356  
Db 945 RMLMEQKEEAMLVATKHATTIDQOQTRRISVLGSDVEKLTAGIABRESSINALSNTWEL 1004  
Qy 357 SLVLEQIEAHATVKEEAFSRKDILEKVERWQNAE--EEAWLEDYNK-----DNRNNA 410  
Db 1005 ISKLETTAELEKJDEL---AVMLKQNSLKNKGEKSEKNEERKKIQDLADQLREAN 1061  
Qy 411 RGAH-----LT----- 416  
Db 1062 KVVHNRMRKNVNLBKKNELDNQVTDLTNKRQLEIQLMDKAAKNEVSGDLLRKMEHDAQ 1121  
Qy 417 -LKA-----EKAR-TLVNKIPGMVDVLTAKAAM--KNERCKEDFTYDGVSL 461  
Db 1122 SMLKQAEQEFRLTLEKVRKALODENORLVNDLATVKAAPFEVKRETSK-----AIS 1174  
Qy 462 SMLDEYMFVQBEKEQKQRQDKKLQDQ-----LKAQEALYGSKPSKPLSTPKA-- 514  
Db 1175 DILDY---RSABEAKANGELDNQRLSDLATVTLKLRQEL--KAKSDNRLRDSQKPE 1230  
Qy 515 -----PHSMGGANRL-----SLGGATMQPPKTTDILHKSVR 547  
Db 1231 EVQSKLANLQKSAVESLQNPMSNRQNSRYVDIPRAASSIGLNMENSDEVLRRSSPSVR 1290  
Qy 548 RAKKTEE---IGTSLSPSSRGLDIAGLP--IKKLSFNASTLRET-----ETPRKPPAQI 596  
Db 1291 FADSSQNRQRAVDSMDVSSGVTVLFLKERIEQLEADNADLSDALEKAKDELQRNEKL 1350  
Qy 597 TFGNSVSTPVRPITNTEDD---ENRTPKFTALNPKTPTMTVAPMQAMT----- 645  
Db 1351 ADQRMVIERVERQLVHITEENTIENR-----MTSQRMVLTNEESSRSR 1395  
Qy 646 -----PSLANKVSAATPVSLVYDKPE 665  
Db 1396 EHEIRSMKARISTLEHLRKEKSLAHLRKEIEVLHGQLHDALESKEKATGLVGVQDSKH 1455  
Qy 666 VTLEQDIDYSPFEERLAI 683  
Db 1456 RDLEEQDRANERELAI 1473

RESULT 10  
TI1722  
hypothetical protein DKFZp434B0435.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: TI1722; T46451; A36881  
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A:Reference number: Z18723  
A:Accession: TI1722  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1780 <POU>  
A:Cross-references: UNIPROT:O9UFR5; EMBL:AL117496  
A:Experimental source: adult testis; clone DKFZp434B0435  
R:Angorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23028  
A:Accession: T46451  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 575-1136, 'C', 1138-1608, 'K', 1610-1780 <AAA>  
A:Cross-references: EMBL:AL137392  
A:Experimental source: adult testis; clone DKFZp434I152  
R:Westendorf, J.M.; Rao, P.N.; Gerace, L.  
Proc. Natl. Acad. Sci. U.S.A. 91, 714-718, 1994  
A:Title: Cloning of cDNAs for M-phase phosphoproteins recognized by the MPM2 monoclonal  
A:Reference number: A36881; MUID:94119956; PMID:8290587  
A:Accession: A36881  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1215-1261, 'E', 1263-1418, 'S', 1420-1608, 'K', 1610-1780 <WES>  
A:Cross-references: GB:L16782; NID:G292328; PIDN:AAC37542.1; PID:G292329  
C:Genetics:  
A:Note: DKFZp434B0435.1; DKFZp434I152.1  
C:Keywords: phosphoprotein

Query Match 6.3%; Score 221; DB 2; Length 1780;  
Best Local Similarity 20.98; Pred. No. 0.0095;  
Matches 164; Conservative 124; Mismatches 305; Indels 190; Gaps 33;  
Qy 1 MSSAVKQDLQHMSTTCDSLLELNLVIMDEVGE---PDTTRDRMLLEQECLEVYRRKV- 56  
Db 995 VSKQVKYRIQEPNRENSPHSSIEAIWEECKEIVKASSKSHQIEELEQQ-IEKLAQAEVK 1053  
Qy 57 ---DOANRSARQAKAIAEAGEAELAGICSANGEPVHVVRQSNQKLHGLRELNIAVYLE 113  
Db 1054 GYKDENNRLK-----EKE-----HKNQDDL-----LKEKETLIQQLKE 1086  
Qy 114 EMKKKKYERNQFVHVIEQIKKISSEIRPADFVPFVQSDLSLRKLD-----ELTKD 168  
Db 1087 ELQEKNTLDVQHQHVEGKRAISELTQGTVCYAKIKELETILETQKVERSHSAKLEQD 1146  
Qy 169 LESIQKE---KSDR-LKQVIEHLNLSLCEVLGIDFKQTVYVHPSLDEAGSKNLSN 223  
Db 1147 I--LEKESIILKLERNLKEFOEHLQ-----DSVKNKTKOLNV 1180  
Qy 224 TTI---ERLAAANRLBEMKIQRKQKLODPASSMELWNLMDTPLEEQMFONITCNIAA 280  
Db 1181 KELKKEEITQLTNNLQDMKHLQLKEEBEETNRQTEKLEBELSASSARTQNLKADLQR 1240  
Qy 281 SEQBITPNTLSDF---LNVSEVLR---EOLKASKMKDLVKK---KAELEEHRRR 331  
Db 1241 KEEDYADLKEKLTDAKQIKQVQKESVVRDEDEKLLRIKINELEKKKNQCSQELDMKQRT 1300  
Qy 332 AHLVGE--GYAEFFSIEAIEAGADPSLVLEQIEAHATVKEEAFSRKD-----ILE 382  
Db 1301 IQQLKEQLNNQKVEEAIQOYERACKDLNVKEKIIDMRMTLEEQEQTVQEQDQVLEAKLE 1360  
Qy 383 KVERWQNACEEBAWLEDYN---KDDNRYNAGRAHLLTKRAEKARTLVNKKIPGMVDVLR 438  
Db 1361 EVERL--ATELEKWKCEKNDLETKNQNSN-----KEHENNTDVLGKLTNLQDELQ 1409  
Qy 439 TKIAAWKNRGEKEDFTYDGVSLSSMLDEYMFVQBEKEKQROROKKLOD-----OLK 492  
Db 1410 ESEQYNADRK-----WLEKMWLITQAKEAENIRNEMKMYADERRFPKQ 1458  
Qy 493 ABEQALYGSKPSKPLSTTKAPRHSNGGA-----NRRLSLGATWQPP 536  
Db 1459 NEMEILTALQLTEDSDLDQKWEERDQVAALEIQLKALISSNVQKONEIQLKRIISETS 1518  
Qy 537 K--TDILHKSVRRAK-----KTEEIGTSLSPSSRGLDIAGLPKLSFNASTLRETET 588  
Db 1519 KIETQIMDIKPKRISSADPKLQTEPLSTSEISRNKIDGVSVDLSCVSTENDQSTRF 1578  
Qy 589 PRKP-----FAQITPGNSVSTP--VRPIT-----NNTEDD-----ENRTPK 623  
Db 1579 P-KPELEIQITPLQPNKMAVKHFGCTTPTVVEIPKARKRKSNEEEDLVCKENKKNATPR 1637  
Qy 624 TTTALNPKTPM-----TVTAPMQAMTP-----SLANKVSAATPVSLVYDKPEVTLEQD 671





Db 362 -RPKMDRMGQIKAKLEKULTPDKVEMVDLLSKAKKEBEKTEKLLKLIANKSSLKTRG 420  
Qy 259 NLMDPTLEBQFQNTCHIAASEQEITEPNTLSDFD---LNVSEVLRQL----- 309  
Db 421 AQLKAVELKSAER-TCVCGRELDDEHRKIMABYETREMKRIAEELAKADEIKKLUKE 479  
Qy 310 -----KASKMKDLVLKKK-----AELEHRR-RAHLVGEEGY 340  
Db 480 RLEKVEKALEKETVLKIQMWDELKALENELSSHDAEKLAESEBYRVKVERLDGLRQ 539  
Qy 341 ABEFSIEATEAGADPSL-----VLEQIEAHIA-----TVKEAFSKOILEKVERWQAC 391  
Db 540 QKILLSSASRIKELSSLEBEALKWVESERGELHRKIREGF---ESLEBLER----- 591  
Qy 392 EEEAMLEDYNDKDNRYNAGRGHILTKRAEKARTLVNKPQMDVLTAKIAAKNERGKE 451  
Db 592 EVQSLRPFYNKMLELKDASRESLELKRREK-----LEDEISEIAKLEANGKA 641  
Qy 452 DFTYDGVSLSSMDLEMYFVRQEKQKQRQDKQLQD--QLKAEQEAALYGGKPSKPL 509  
Db 642 E-----ETRGQIDELLRIYSBEHRLSDEHLRKSKELAGLKSRLTRESLSQSAEKOL 695  
RESULT 15  
MWKW  
Myosin heavy chain B [similarity] - Caenorhabditis elegans  
N;Contains: myosin ATPase (EC 3.6.4.1)  
C;Species: Caenorhabditis elegans  
C;Date: 13-Jun-1983 #sequence revision 19-May-2000 #text change 09-Jul-2004  
C;Accession: T20770; T21629; A93958; A93287; A21074; A02992  
R;Kershaw, J.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19322  
A;Accession: T20770  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1963 <WIL>  
A;Cross-references: UNIPROT:O02244; EMBL:Z81499; PIDN:CAB04089.1; GSPDB:GN00019; CESP:F1  
A;Experimental source: Clone F1C3  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1963 <W12>  
A;Cross-references: EMBL:Z83107; PIDN:CAB05505.1; GSPDB:GN00019; CESP:F11C3.3  
A;Experimental source: Clone F32A7  
R;Karn, J.; Brenner, S.; Barnett, L.  
Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983  
A;Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy ch  
A;Reference number: A93958; MUID:83273600; PMID:6576334  
A;Accession: A93958  
A;Molecule type: DNA  
A;Residues: 1-61, 'EMSVIQ', 65-376, 'V', 378-1963 <KAR>  
A;Cross-references: GB:J01050; NID:g156399; PIDN:AAA28124.1; PID:g156400  
R;McLachlan, A.D.; Karn, J.  
Nature 299, 226-231, 1982  
A;Title: Periodic charge distributions in the myosin rod amino acid sequence match cross  
A;Reference number: A93287; MUID:82272395; PMID:7202124  
A;Accession: A93287  
A;Molecule type: DNA  
A;Residues: 847-1333, 'R', 1335-1876, 'L', 1878-1963 <MCL>  
R;Wills, N.; Gesteland, R.F.; Karn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.  
Cell 33, 575-583, 1983  
A;Title: The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonser  
A;Reference number: A21074; MUID:8322892; PMID:6571695  
A;Accession: A21074  
A;Molecule type: DNA  
A;Residues: 1873-1963 <W13>  
A;Cross-references: GB:V01494; GB:J01049; NID:g6783; PIDN:CAA24738.1; PID:g6784  
C;Genetics:  
A;Gene: unc-54; CESP:F11C3.3  
A;Map position: 1  
A;Introns: 21/3; 64/3; 111/3; 264/1; 525/3; 951/2; 1747/3; 1819/3; 1894/3

C;Superfamily: myosin heavy chain; myosin motor domain homology  
C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle c  
F;84-775/Domain: myosin motor domain homology <MMOT>  
F;174-181/Region: nucleotide-binding motif A (P-loop)  
F;662-684/Region: actin binding #status predicted  
F;766-780/Region: actin binding #status predicted  
F;848-1963/Domain: coiled coil #status predicted <COI>  
F;848-1162/Region: S2  
F;1163-1963/Region: light meromyosin  
F;125/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted  
F;180/Binding site: ATP (Lys) #status predicted  
F;702,712/Active site: Cys #status predicted  
Query Match 6.1%; Score 213.5; DB 1; Length 1963;  
Best Local Similarity 20.6%; Pred. No. 0.024;  
Matches 150; Conservative 125; Mismatches 295; Indels 157; Gaps 27;  
Qy 1 MSSAVKQDLQHQMSSTTCDSLLLELVINDEVGEPTDTRMLLEBOECLEEVYRRKVDQAN 60  
Db 1313 LKSQTSQLEEARRTADEEARERQTVAQAQKNYQHEAEQLESLEE-IEGNEILRQLS 1371  
Qy 61 RSRAQLR--KATAEGEAEAGICSAMGPPVHVHVRQSNQKLHGLREELNAIVPYLEBMKK 118  
Db 1372 KANADIQWKARFEGE---GLLKA-DELEDAKRQAQKINELQEAALDAANSKNASLEKT 1426  
Qy 119 K-----VERWNQFVHVIEQ---IKKISSEIRPADFPVFKVPVQSDLSLRK 151  
Db 1427 KSLRGVGLDDAQVDVERANGVASALEKKQKQKPKDIIDWR-----KK 1468  
Qy 162 LDELTKDLESLOKESDRLLKQVIEHLSLHSLCEVL-GI-----DFKQTVYVHPSLDEA 215  
Db 1469 TDDLAELDGAQDRDLRTSTDLFKAKNAQAEELAEVVEGVRRENKSLSQIKOLTDQLEG 1528  
Qy 216 EGSKNLSNTTIERLAAANRLREMKIQRMQKLDQFASSMLEL--WNLMPTPEEQMQFN 273  
Db 1529 GRSVHEMKIIRRL-----EIEKELOHALDEAAEALESKVLRAQAEVSVIRSE 1580  
Qy 274 ITCNTAASQETETPNTLSTDFLNVSSEVLRLQKAS-----KMKDLVLKKA--- 323  
Db 1581 IEKRIQKEEE-----PENTRNHARALESMQASLETEAKGKAEALLRIKKKLEGD 1630  
Qy 324 ---ELE---EHRRAHLVGEEGVAEFSIEAIEAGALDPSLVLEQIEAHATVKEEAFS-- 376  
Db 1631 INELEIALDHANKADAKQ-----NLKRYEQVRELQVLEEQVEEORNGADTRQEFNAE 1684  
Qy 377 -----RKDIL---EKVERWQNAEEAWLEDYNDKDNRYNAGRGHILTKRAEKAR 424  
Db 1685 KRATLQSEKEELLVANAEEARQAQAEYA--ADARQANANAAQVSSILTSAKR----- 1737  
Qy 425 TLVNKIPGMVDVLRTKIAAAWKNR-RGKEDFTYDGVSLSSMLDEYMFVRQEKQKQRD 483  
Db 1738 ---KLEGEIQAIHADLDLTLNEYKAAERSKKAATADATRLAEL--ROEQEHSQHVDR 1791  
Qy 484 QKKLQDLK-----AQEALYGSKPSKPLSTTKAPRHSMGGANRL-----SLGG 530  
Db 1792 RKGLEQQLKEIQVRLDEAEAAALKGGKVIKLEQRVRELESELDEGEQRFPQAKNLG- 1850  
Qy 531 ATMQPPTDILHSKSVRAAKTKEEIGTSLSPSSRGDIDAGLPKIKLSFNASTLRETETPR 590  
Db 1851 -----RADRRVRELQFQVDEKKNFERLQDLIDKLQ-----QKLTQK 1888  
Qy 591 KPFAQITPGNSVSTPVRPITNNEDENRTPKTFTALNPKTPTMTVTPMQMWTSLAN 650  
Db 1889 KQVEEAELANLNQYKQLTHQLEDAERADQ---AENSLSKVRSKRSASAVAPQLOS 1945  
Qy 651 KVSATPV 657  
Db 1946 SASAAVI 1952  
Search completed: November 20, 2004, 00:42:40  
Job time : 31 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 19:16:05 ; Search time 87 Seconds

(without alignments)  
2845.094 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKQDLHQHSTTCDSLL.....DIDYFERRLAIVLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep 23Sep04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3490	99.8	690	5	ABB05052 Rice leaf
2	3477.5	99.5	690	5	Aau97649 Plant sha
3	1236.5	35.4	578	3	Aag20828 Arabidops
4	756.5	21.6	338	3	Aag20829 Arabidops
5	693.5	19.8	318	3	Aag20830 Arabidops
6	325.5	9.3	620	6	Abu70658 Human adi
7	323.5	9.3	620	8	Adn04802 Antipsori
8	323.5	9.3	620	8	Ado20037 Human PRO
9	323.5	9.3	620	8	Adq09246 Human PRC
10	319.5	9.1	657	4	Abb11530 Human cyt
11	313.5	9.0	670	7	Adb78256 Human CGD
12	267.5	7.7	639	8	Adp04701 Sea squir
13	240	6.9	1857	5	Aau84350 Protein M
14	240	6.9	1938	6	Abn92127 Human cer
15	240	6.9	1945	6	Abu10398 Human smo
16	240	6.9	1972	6	Abn92126 Human cer
17	240	6.9	1972	7	Adc58724 Human Pro
18	240	6.9	1972	7	Adc58721 Human Pro
19	240	6.9	1972	7	Adc63520 Human Pro
20	240	6.9	1972	7	Adc63516 Human Pro
21	240	6.9	1972	8	Adk70527 Respirato
22	240	6.9	1972	8	Adn03745 Antipsori
23	240	6.9	1979	6	Abu10399 Human smo
24	237.5	6.8	953	4	Abu53071 Intracell
25	237.5	6.8	2228	7	Abn61599 Human gol

26	237.5	6.8	2230	6	ABU07445	Abu07445 Protein d
27	237.5	6.8	2230	7	ABR61600	Abn61600 Human gol
28	237.5	6.8	2250	7	ABR61601	Abn61601 Human gol
29	237.5	6.8	2252	7	ABR61602	Abn61602 Human gol
30	237	6.8	941	4	ABU53072	Abu53072 Intracell
31	237	6.8	953	4	ABU53069	Abu53069 Intracell
32	235	6.7	860	5	AD117302	Ad117302 Polypepti
33	235	6.7	860	7	ADC38517	Adc38517 Human AML
34	233.5	6.7	1455	7	ADe80783	Adc80783 Microsate
35	230	6.6	880	4	AA896332	Aab96332 Putative
36	229	6.6	885	2	AA866930	Aar66930 AML chro
37	224.5	6.4	907	4	ABU53073	Abu53073 Intracell
38	221.5	6.3	876	7	ADM26084	Adm26084 Hyperther
39	221.5	6.3	951	4	ABU53070	Abu53070 Intracell
40	221	6.3	1114	7	ADJ70313	Adj70313 Human hea
41	221	6.3	1213	4	AA440016	Aam40016 Human pol
42	221	6.3	1213	6	ABP97031	Abp97031 Human L-F
43	221	6.3	1780	4	AA38681	Aam38681 Human pol
44	221	6.3	1780	4	ABU53201	Abu53201 Human cel
45	221	6.3	1879	4	AA25750	Aam25750 Human pro

ALIGNMENTS

RESULT 1

ABB05052

ID ABB05052 standard; protein; 690 AA.

XX ABB05052;

XX DT 26-MAR-2002 (first entry)

XX DE Rice leaf shape and size controlling protein SEQ ID NO:2.

XX KW Rice; leaf shape and size controlling gene; leaf shape; leaf size.

XX OS Oryza sativa.

XX FH Key Location/Qualifiers

FT Misc-difference 170

FT /label= unknown

FT /note= "encoded by GNA"

XX PN JP2001258574-A.

XX PD 25-SEP-2001.

XX PF 23-MAR-2000; 2000JP-00083067.

XX PR 23-MAR-2000; 2000JP-00083067.

XX XX (NORQ ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.

PA (SEIB-) SEIBUTSUKETI TOKUTEI SANGYO GIJUTSU.

PA (NORI-) ZH NORIN SUISAN SENTAN GIJUTSU SANGYO.

XX WPI: 2002-037587/05.

DR N-PSDB; ABA92702.

DR A new gene controlling leaf shape.

XX PS Claim 1; Page 9-11; 14pp; Japanese.

XX CC The present sequence represents the protein encoded by a polynucleotide

CC (see ABA92702) encoding a plant gene which can control leaf shape and

XX size. The polynucleotide can be used for controlling leaf shape

XX SQ Sequence 690 AA;

Query Match 99.8%; Score 3490; DB 5; Length 690;

Best Local Similarity 99.9%; Pred. No. 1.7e-245;

Matches 689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



```
Db      600 SVSSTVTRITNNTDENRTKFTALNPKPTMTVTAPMQMTPSLANKVTSATPVSIV 659
Oy      661 YDKPEVTLOEDIDYSEERRLAIYLARQMV 690
Db      660 YDKPEVTLOEDIDYSEERRLAIYLARQMV 689

RESULT 3
AAG20828
ID AAG20828 standard; protein; 578 AA.
AC
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23163.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2..
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
```



PR 28-APR-1999;	99US-0131449P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145218P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134218P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134219P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145918P.
PR 14-MAY-1999;	99US-0134370P.	PR 28-JUL-1999;	99US-0145919P.
PR 18-MAY-1999;	99US-0134768P.	PR 28-JUL-1999;	99US-0145951P.
PR 19-MAY-1999;	99US-0134941P.	PR 02-AUG-1999;	99US-0146386P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146388P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146389P.
PR 24-MAY-1999;	99US-0135629P.	PR 03-AUG-1999;	99US-0147038P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147204P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147302P.
PR 28-MAY-1999;	99US-0136782P.	PR 05-AUG-1999;	99US-0147192P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147260P.
PR 03-JUN-1999;	99US-0137528P.	PR 06-AUG-1999;	99US-0147303P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147416P.
PR 07-JUN-1999;	99US-0137724P.	PR 09-AUG-1999;	99US-0147933P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138840P.	PR 10-AUG-1999;	99US-0148171P.
PR 10-JUN-1999;	99US-0138847P.	PR 11-AUG-1999;	99US-0148311P.
PR 14-JUN-1999;	99US-0139119P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139452P.	PR 13-AUG-1999;	99US-0148565P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148684P.
PR 17-JUN-1999;	99US-0139492P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139454P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139455P.	PR 18-AUG-1999;	99US-0149428P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149903P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 14-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144332P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.

PR	18-OCT-1999;	99US-0159584P.	PR	05-MAR-1999;	99US-01231800P.
PR	21-OCT-1999;	99US-0160741P.	PR	09-MAR-1999;	99US-0123548P.
PR	21-OCT-1999;	99US-0160767P.	PR	23-MAR-1999;	99US-0125788P.
PR	21-OCT-1999;	99US-0160768P.	PR	25-MAR-1999;	99US-0126264P.
PR	21-OCT-1999;	99US-0160770P.	PR	29-MAR-1999;	99US-0126785P.
PR	21-OCT-1999;	99US-0160814P.	PR	01-APR-1999;	99US-0127462P.
PR	21-OCT-1999;	99US-0160815P.	PR	06-APR-1999;	99US-0128234P.
PR	22-OCT-1999;	99US-0160980P.	PR	08-APR-1999;	99US-0128714P.
PR	22-OCT-1999;	99US-0160981P.	PR	16-APR-1999;	99US-0129845P.
PR	22-OCT-1999;	99US-0160989P.	PR	19-APR-1999;	99US-0130077P.
PR	25-OCT-1999;	99US-0161404P.	PR	21-APR-1999;	99US-0130449P.
PR	25-OCT-1999;	99US-0161405P.	PR	23-APR-1999;	99US-0130510P.
PR	25-OCT-1999;	99US-0161406P.	PR	28-APR-1999;	99US-0130891P.
PR	26-OCT-1999;	99US-0161359P.	PR	30-APR-1999;	99US-0131449P.
PR	26-OCT-1999;	99US-0161360P.	PR	30-APR-1999;	99US-0132048P.
PR	26-OCT-1999;	99US-0161361P.	PR	04-MAY-1999;	99US-0132407P.
PR	28-OCT-1999;	99US-0161920P.	PR	05-MAY-1999;	99US-0132484P.
PR	28-OCT-1999;	99US-0161922P.	PR	06-MAY-1999;	99US-0132485P.
PR	28-OCT-1999;	99US-0161993P.	PR	06-MAY-1999;	99US-0132486P.
PR	29-OCT-1999;	99US-0162142P.	PR	06-MAY-1999;	99US-0132487P.
Query Match 21.6%; Score 756.5; DB 3; Length 338;					
Best Local Similarity 45.6%; Pred. No. 1.3e-46;					
Matches 154; Conservative 69; Mismatches 104; Indels 11; Gaps 4;					
QY	241	IORMKIQDFASSMLELNWMLDTPLEEQOMFONITCNIAASEQBITPTLTSTDFLNYVE	300	PR	14-MAY-1999;
Db	1	MORLKKQLATQUTDWNMLDTSDEERELFHVTSNISASVHEVTASGALALDLIEQAE	60	PR	14-MAY-1999;
QY	301	SEVLRLQKASKMKOLVLKKKAELEHRRRAHLVGEEGYAEFEFSIBAIAGAIIDPSLVL	360	PR	14-MAY-1999;
Db	61	VEVDRLDQLKSSRMKEIAFKQSELEIYARAHIEIKPEVVREIRIMSLIDAGNTEPTELL	120	PR	14-MAY-1999;
QY	361	EQIEAHIAIVKEAFSRKIDILEKVERWQACEEAWLEDYKDDNRYNAGRGHILTKRA	420	PR	18-MAY-1999;
Db	121	ADMSQIAKAEAFSRKEILDRVKKWSACEESWLEDYKDDNRYNAGRGHILTKRA	180	PR	18-MAY-1999;
QY	421	EXARTLVNKPQWVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVROEKQEKKR	480	PR	19-MAY-1999;
Db	181	EXARILVSKITAMVDTLAKTAWEEENSMSFEYDGVPLLAMLDYETMLRQERDEKRR	239	PR	20-MAY-1999;
QY	481	QRDQKKLQDLKAEQALYGVKSPSKPLSTKKAPRHSWGGANRLSLGGATMQPPKTDI	540	PR	21-MAY-1999;
Db	240	LKEQKQKQEPHTDQESAFGSKFSPARVPSAKP-----VGTRVGGGLNETPMRRLS	292	PR	24-MAY-1999;
QY	541	LHKSVAARKTEETIGTL-SPSS--SRGLDIAGLPIKK	575	PR	25-MAY-1999;
Db	293	MNSNQNGSKSRSLNKIASPSNIVANTKDDAASPVR	330	PR	27-MAY-1999;
RESULT 5					
AAG20830					
ID	AAG20830 standard; protein; 318 AA.				
XX					
AC	AAG20830;				
XX					
DT	17-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 23165.				
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; Genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-00301439.				
XX					
PR	25-FEB-1999; 99US-0121825P.				

PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149829P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151910P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155513P.
PR	23-SEP-1999;	99US-0155548P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 19.8%; Score 693.5; DB 3; Length 318;		
Best Local Similarity 45.3%; Pred. No. 4.8e-42;		
Matches 144; Conservative 62; Mismatches 101; Indels 11; Gaps 4;		
QY	261	MDTPLEEQMFONITCNIAASRQEITEPTNTLTSTDFLNYVESEVLRLEQLKASKMKDLVLK 320
Db	1	MDTSDEERELFDHVTNSISASVHVTASGALALDLIEQAEVEVDRLDQLKSSRMKEIAFK 60
QY	321	KKAELERHRRRAHLYGEEGYAEFSEIAEAGAI DPSLVLEQIEAHIAIATVKEEAFSRKDI 380
Db	61	KQSELEETIYARAHIEIKPEVVRERIMSLIDAGNTEPTTELLADMDSDQIAKAKEEAFSRKEI 120
QY	381	LEKVERWQACEEAEWLEDYKDDNRVYAGRAHLTLKBAEKARTLVNKI PGWVDVLRTK 440
Db	121	LDRVEKMSACSEESWLEDYKDDNRVYAGRAHLTLKBAEKARTLVNKI PGWVDVLRTK 180
QY	441	IAAKWNERGKEDFTDGVSLSSMLDEYMFVROEKEQKQRDQKQLDQKAEQBALYG 500
Db	181	TRAWEEENSM-SFEYDGVPLLAHLDEYTLWROEREDEKRLKEQKQEQPHPTDQSSAFG 239
QY	501	SKPSPSKPLSTKKAPRHSMGGANRLSLGGATWQPPKTDILHKSVRAAKKEEIGTL-S 559
Db	240	SKPSPARPVSAKKP-----VGTRVNGGGLNETPMRRLSMNSNQNGSKSRDLSNKIAS 292
QY	560	PSS--SRGLDIAGLPKK 575
Db	293	PSNIVANTKDDAASPVS 310
RESULT 6		
ABU70658		
ID ABU70658 standard; protein; 620 AA.		
XX		
AC ABU70658;		
XX		
DT 10-JUN-2003 (first entry)		
XX Human adipocyte Selected Interacting domain, SID, #289.		
DE		
XX		





PD	27-MAY-2004.				
XX					
PF	06-NOV-2003; 2003WO-US035268.				
XX					
PR	08-NOV-2002; 2002US-0425235P.				
XX	(GETH ) GENENTECH INC.				
XX					
PI	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;				
PI	Wood WI, Wu TD;				
XX					
XX	WPI: 2004-420067/39.				
DR	N-PSDB; ADO20036.				
XX					
PT	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for				
PT	treating an immune related disorder such as systemic lupus erythematosus, or				
PT	rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or				
PT	spondyloarthropathy.				
XX					
PS	Claim 7; SEQ ID NO 946; 173lpp; English.				
XX					
CC	The invention relates to human PRO polypeptides and the polynucleotides				
CC	encoding them. The polypeptides and polynucleotides are useful for				
CC	treating and diagnosing immune related disorders in mammals. The immune				
CC	related disorders include systemic lupus erythematosus, rheumatoid				
CC	arthritis, osteoarthritis, juvenile chronic arthritis, systemic				
CC	sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune				
CC	haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes				
CC	mellitus, immune-mediated renal disease, demyelinating diseases of the				
CC	central or peripheral nervous system, demyelinating polyneuropathy,				
CC	Guillain-Barre syndrome and chronic inflammatory demyelinating				
CC	polyneuropathy. This sequence represents a human PRO polypeptide of the				
CC	invention.				
XX					
SQ	Sequence 620 AA;				
Query Match		9.3%;	Score 323.5;	DB 8;	Length 620;
Best Local Similarity		22.7%;	Pred. No. 1.2e-14;		
Matches 144;		Conservative 124;	Mismatches 266;	Indels 101;	Gaps 24;
QY	9 LHMSTTC-DSLLLELNVIWDEVGEPDTRDRMLLELEQECLEVVYRRKVDQANRSRAQLR	67			
Db	7 LAEESIVCLQKALNHLREIWEIGIPEDQRLQRTVEVKKHIELLDMMIAEBSLERLI	66			
QY	68 KAIAGEARLAGICSNAGPVPVHVRQSNQKLGHLREELNAIVPYLEMKKKKVWQNVQFV	127			
Db	67 KSIQVCQKELNLTCLSELHVEPQ-EEGETTILQLEKDLRTQV---ELMRKQKKEKQELK	122			
QY	128 HVIEQIKTSSIRPADFPVFKVPVDQSDL-SLRKLDLTKLESLOKESDRLKQVI--	184			
Db	123 LLQEQDQELCEILCMPHY-----DIDSASVPSLEELNQPHQVTTLTRETKASRREBFSVI	177			
QY	185 --EHLNLSHSLCEVLGIDFKQTVYVHPSLDRAEGSKNLSNTTIERLAAANRLRMKIQ	242			
Db	178 KRQIILCMBEELDHTPDTSFERDV-----VCEDEDAFCLS---LENIATLQKLLRQLEMQ	228			
QY	243 RMQK---LQDFASSMLELNLMDTPLEEQMFQNTICNTAAEQEITEPTNLSTDFLNVV	299			
Db	229 KSNQNEAVCEGLRTQIRELWDLRQIPEEREAVATI---MSGKAKVRKA-----L	275			
QY	300 ESEVLRLQELKASKMDLVKKKAEI-----EHRRAHLVGEEGYAEFSEIAI	349			
Db	276 QLEVDRLLEELKQNMKKVIEAIRVELVQWDQCFYSQEQRQAFAPCAEDYTES-----	329			
QY	350 EAGAI DPSVLVLEQIEAHIAITVKEEAFSRKDI LEKVERWQNVACBEEAW-----LEDYKDD	404			
Db	330 -----LLQLHDAEIVRLKNYVEVHKELFEGVQKW-----EETWRLFLPERKASDP	375			
QY	405 NRYNAGRGHILTKRAEKARTLVNKPVGMDVLRTKIAAWKNERGKEDFTYDGVSLSSML	464			
Db	376 NRF-TNRGGNL-LKEEKQRAKQKMLPKLEBELKARIELWEQHSKA-FWVNGQKFM---	429			
QY	465 DEYMFVROKEQ-EKKRQDQKKLODLKAEQALYGSKP-----SPSKPLSTKKA	514			
Db	430 -EYVAEQWEMHLEKRAQERQKLNKQKTEEMLYGSAPRTPSKRGGLAPNTPGKARKL	488			
QY	515 PRHSMGGANRLSL-----GGATMQPPKTDILHSKSVRAAKTEETIGLSPSSRGLDIAG	570			
Db	489 NTTTMSNATANSIRPIFGTVYHSPVSRLLPPSGSKPVAASTCS-GKKTPTGR-----HG	543			
QY	571 LPKIKLSFNASTLRETETPRKPAQTTPGNSVST	605			
Db	544 ANKENLENGSILSGYGPAGSAPLQRNFSINSVAST	578			
RESULT 8					
ADO20037					
ID	ADO20037 standard; protein; 620 AA.				
XX					
AC	ADO20037;				
XX					
DT	12-AUG-2004 (first entry)				
XX					
DE	Human PRO polypeptide #473.				
XX					
KW	Human; PRO; immune related disorder; systemic lupus erythematosus;				
KW	rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;				
KW	systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;				
KW	autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;				
KW	diabetes mellitus; renal disease; demyelinating disease;				
KW	central nervous system; peripheral nervous system;				
KW	demyelinating polyneuropathy; Guillain-Barre syndrome;				
KW	chronic inflammatory demyelinating polyneuropathy.				
XX					
XX	Homo sapiens.				
XX					
PN	WO2004043361-A2.				
XX					

```

Db 430 -EYVAEQWEMHRLERAKQERQKXKQKTEMLYGSAPRTSPSKRGLAPNTPCKARKL 488
QY 515 PRHSWGGANRRSL-----GGATMQPPKTDILHKSVAARKTEETGTLSPPSSRGLDIAG 570
Db 489 NTTTMSNATANGSIRPIFGTGYHSPVRLPPSGKPKVAASCTCS-GKKTPTGR----HG 543
QY 571 LPIKLSFNASTLRETEPRKPFQAITPNSVSST 605
Db 544 ANKENLELNGSILSGYPGSAPLQRFNSINSVAST 578

RESULT 9
ADQ09246
ID ADQ09246 standard; protein; 620 AA.
XX
AC ADQ09246;
XX
23-SEP-2004 (first entry)
XX
Human PRC1 protein SEQ ID NO:431.
DE
XX
KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human.
XX
OS Homo sapiens.
XX
PN WO2004055050-A2.
XX
PD 01-JUL-2004.
XX
PF 10-DEC-2003; 2003WO-IB006434.
XX
PR 10-DEC-2002; 2002US-0432699P.
XX
PR 03-JUL-2003; 2003US-0485027P.
XX
PA (ENDO-) ENDOCYTE SAS.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Girard J, Amalric F, Roussigne M, Clouaire T;
XX
WPI; 2004-525034/50.
XX
DR N-PSDB; ADQ09247.
XX
Modulating expression of a Thanatos (death)-Associated Protein (THAP)
PT responsive gene for preventing or treating e.g. cancer or inflammation,
PT comprises modulating the interaction of a THAP polypeptide with a nucleic
PT acid.
XX
Example 47; SEQ ID NO 431; 612pp; English.
XX
The present invention describes a method for modulating the expression of
CC a thanatos (death)-associated protein (THAP) responsive gene. The method
CC comprises modulating the interaction of a THAP-family polypeptide or its
CC biological fragment with a nucleic acid, and so enhancing or repressing
CC the expression of the THAP responsive gene. Also described: (1) a method
CC of modulating the expression of a gene responsive to a THAP/chemokine
CC complex; (2) a pharmaceutical composition comprising a THAP responsive
CC element in a pharmaceutical carrier; (3) a transcription factor decoy
CC consisting essentially of a THAP responsive element; (4) a cell
CC comprising a transcription factor decoy described above; (5) methods of
CC modulating the interaction between a nucleic acid and a THAP-family
CC polypeptide or its biological fragment, or a nucleic acid and a
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC cell comprising a viral vector which comprises a promoter operably linked
CC to a nucleic acid encoding a THAP-family polypeptide or its biological
CC fragment; (7) a method of constructing a cell which expresses a
CC recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex

```

```

CC ; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC comprising a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation us useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 620 AA;

```

```

Query Match          9.3%; Score 323.5; DB 8; Length 620;
Best Local Similarity 22.7%; Pred. No. 1.2e-14;
Matches 144; Conservative 124; Mismatches 266; Indels 101; Gaps 24;

QY 9 LHMSTTC-DLLELVINWDEVEGPDTRDRMLLEQECLEVYRRKVDQANRSRQLR 67
Db 7 LAESIVCLOKALNHLREIWEIGIPEDQRLQRTVEVYKGIKELLDWMAIEESLKERLI 66
QY 68 KAIAGEAEAGAGICSAMGEPPVHVRSQKQLHGLREELNAIVPYLEEMKKKVKRQNFV 127
Db 67 KSIVCQKELNTLCSLHVFPFO-EGETTLQLEKDLRTQV---ELMRKQKBERKELK 122
QY 128 HVIEQIKKISSEIRPADVPFPKVPVQSDL-SLRKLDLTKDLSLOKESDRLKQVI-- 184
Db 123 LLQEQDQELCELCPHY----DIDSASVPSLEELNQFRQHVTTLRETKASREEFVSI 177
QY 185 --EHLNLSLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNTTIERLAAAANRLRMKIQ 242
Db 178 KRQIILCMBELDHTPDTSFERDV-----VCDESDAFCLS---LENTATLQKLRLQEMQ 228
QY 243 RMQK---LQDFASSMLLELWMLDTPLEEQMFQNTCNIAASEQEIETEPNTLSTDFLNV 299
Db 229 KSQNEAVCEGLRTOIRELMDRLQIPEEREAVATI---MSGSKAKVKA-----L 275
QY 300 ESEVLRLQKASKMKDVLVKKKAEI-----BEHRRRAHLVGEEGYAEBSFIEAI 349
Db 276 QLEVDRLLEELKMQNKKVIEAIRVELVQYDQCFVSEQRQAPAFPACEDYTES----- 329
QY 350 EAGADISLVLEQIEAHIAIVTKEAFSRKDIILEKVERWQNCAREAW-----LEDYKDD 404
Db 330 -----LQLHDAEIVRLKNYVEHKLFEVQKWM-----BETWRLFLEFPERKASDP 375
QY 405 NRYNAGRGALHTLKRAEKARTLVNKIPQWVDVLRTKIAAWKNRGKEDFTYDGVLSLSSML 464
Db 376 NRF-TNKGGLN-LKEEKQRAKLQKMLPLSEELKARLEWQEHASKA-FWYNGQKFM---- 429
QY 465 DEYMFVRQEKEQ-EKKQRQDKKLQDQKAFQEAALYGSKP-----SPSKPLSTKKA 514
Db 430 -EYVAEQWEMHRLERAKQERQKXKQKTEMLYGSAPRTSPSKRGLAPNTPCKARKL 488
QY 515 PRHSWGGANRRSL-----GGATMQPPKTDILHKSVAARKTEETGTLSPPSSRGLDIAG 570
Db 489 NTTTMSNATANGSIRPIFGTGYHSPVRLPPSGKPKVAASCTCS-GKKTPTGR----HG 543

```



AC	ADD78256;	Best Local Similarity 22.5%; Pred. No. 7e-14;	
XX		Matches 156; Conservative 122; Mismatches 262; Indels 153; Gaps 30;	
DT	29-JAN-2004 (first entry)		
XX			
DE	Human CGDD-37.		
XX			
KW	Anabolic; Hypertensive; Respiratory; Anti-HIV; Antiallergic;		
KW	Neuroprotective; Nootropic; Antianemic; Antiartherosclerotic;		
KW	Antiinflammatory; Ophthalmological; Muscular; Hepatotropic;		
KW	Neuroprotective; Antispasmodic; Anticonvulsant; Virucide; Antibacterial;		
KW	Fungicide; Antiparasitic; Protozoicide; Antihelminthic; Cytostatic;		
KW	Cerebroprotective; Antiparkinsonian; Antipsoriatic; Antigout;		
KW	Antidiabetic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;		
KW	human; cell growth; cell differentiation; cell death; CGDD;		
KW	cell proliferative disorder; cancer; developmental disorder;		
KW	neurological disorder; autoimmune disorder; inflammatory disorder;		
KW	infection; reproductive disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003077875-A2.		
XX			
PD	25-SEP-2003.		
XX			
XX	14-MAR-2003; 2003WO-US008310.		
XX			
PR	15-MAR-2002; 2002US-0364494P.		
PR	29-MAR-2002; 2002US-0369129P.		
PR	12-APR-2002; 2002US-0372511P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Kable AE, Tran UK, Hafalia AJA, Burford N, Honchell CD;		
PI	Lehr-Mason PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;		
PI	Elliot VS, Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;		
PI	Becha SD, Lee SY, Swarnakar A, Yue H, Warren BA, Baughn MR, Lal PG;		
PI	Lee S, Ho A, Gandhi AR, Yao MG;		
XX			
DR	WPI: 2003-779081/73.		
DR	N-PSDB; ADD78295.		
XX			
PT	New polypeptides and polynucleotides associated with cell growth,		
PT	differentiation and death, useful for diagnosing, treating or preventing		
PT	e.g. developmental, neurological, autoimmune, inflammatory or		
PT	reproductive disorders.		
XX			
PS	Claim 1; SEQ ID NO 37; 320pp; English.		
XX			
CC	The present invention relates to novel human proteins (I; ADD78220-		
CC	ADD78256) and their coding sequences (II; ADD78259-ADD78297), which are		
CC	associated with cell growth, differentiation and death, referred to as		
CC	CGDD-n proteins, where n is a number from 1 to 39. The CGDD proteins and		
CC	their coding sequences are useful for diagnosing, treating or preventing		
CC	cell proliferative disorders (e.g. cirrhosis, hepatitis,		
CC	arteriosclerosis, psoriasis, primary thrombocytopenia) or cancers (e.g.		
CC	adenocarcinoma, sarcoma or cancers of the bone, bone marrow, brain,		
CC	breast, colon, kidney, liver, lung or uterus), developmental disorders		
CC	(e.g. renal tubular acidosis, Becker muscular dystrophy, gonadal		
CC	dysgenesis, hypothyroidism or seizures), neurological disorders (e.g.		
CC	Pick's disease, cataract, epilepsy, ischemic cerebrovascular disease,		
CC	stroke, Alzheimer's disease, Parkinson's disease or dementia),		
CC	autoimmune/inflammatory disorders (e.g. AIDS, allergies, anemia, asthma,		
CC	diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid		
CC	arthritis, contact dermatitis or gout), viral, bacterial, fungal,		
CC	parasitic, protozoan or helminthic infections, reproductive disorders		
CC	(e.g. infertility, ectopic pregnancy, premature ovarian failure, delayed		
CC	puberty or prostaticitis) or disorders of the placenta (e.g. preeclampsia,		
CC	choriocarcinoma, placenta previa, placental or maternal floor infarction		
CC	or chronic villitis).		
XX			
SQ	Sequence 670 AA;		

Query Match

9.0%; Score 313.5; DB 7; Length 670;

QY	9 LHQMSTTC-DSLLLELVINWDEVEGPDTRDRMLLELEOECELEVYRRKVDQANRSRAQLR	67
DB	7 LAEESIVCLOKALNHLREINELIGIPEDQRLQRTVEVVKKHIKELDDMMMAEESLKERLI	66
QY	68 KAIAGEAGELAGICSAMGEPVHVQRQSNOKLHGLREELNAIVPYLEEMKKKKVERWNOFV	127
DB	67 KSIISVCQKELNTLSEHVEFQ-RQSLAVLS--RQECNGV--SAHCNURLPGSSDFP	120
QY	128 HVIQIKKISSEIRPADFPVF--KVPVDSQSLSLRKLDELTKDLESLOREKS	177
DB	121 AVASRVAGITG-VRRHAWLIFLVKTGFHVHVEGETTILQLEKDLRTQVLMRKQKK	179
QY	178 DRLQVTEHLNLSL-HSICEVL--GID--FKQTVVEVHPSLEAEQSK	219
DB	180 ER-KQELKLLQEOQELCEILCMPHYDIDSASVPSLELNQFRQHV---TTLRETKASR	234
QY	220 -----NLSNT-----TIERLAAAANRLRMKIQRM	244
DB	235 REEFVSIVKRIILCWEELDHTPDTSPERDVVCEDEDAFCLSLNENIATLQKLLRQLEWQKS	294
QY	245 QK---LQDFASSMLELWNLMDTPLEEQOMFQNTICNIAASQEIETEPNTLSTDFLNYYES	301
DB	295 QNEAVCEGLRTQIRELWDLQIPEEREAVATI--MSGSKAKVKA-----LQL	341
QY	302 EVLRLEQLKASKMDLVKKAEI-----EHRRAHLVGEQYAEFSEIAIEA	351
DB	342 EVDRLEELKQNMKKVIEAIRVELVQWDQCFYSQEQRAFAPFCAEDYTES-----	393
QY	352 GAIDPSIVLEQIEAHIAIVKEEAFSRKIDLEKVERWQACEEAW-----LEDYNKDDNR	406
DB	394 -----LLQHDAAEIVRLKNYYEVHKEPFGVQKW-----EETWRLFEPERKASDNR	441
QY	407 YNAGRGHAHLTKRAEKARTLVNKIPGMVDVLRITKIAAWKNERGKEDFTYDGVLSMLDE	466
DB	442 F-TNRGGNL-LKEBKQRAKLQKMLPKEEELKARIELWEQESHKA-FWNGQKFW-----E	494
QY	467 YMFVROEKEQ-EKKRQDQDKLQDLKAEQALVGSKP-----SPSKPLSTKKAPR	516
DB	495 YVAEQWEMHLEKERAQERQKQKQTEMLYGSAPRTPSKRRGLAPNTPGKARKLNT	554
QY	517 HSMGAGNRRLSL-----GGATWQPPKTDILHKSVRAAKTEETIGTLPSSSRGLDIAGLP	572
DB	555 TTMSNATANSIRPIFGTVVHSPVSRIPPSGSKPVAASICS-GKKTPTGCR-----HGAN	609
QY	573 IKKLSFNASTLRETETPRKPPAQITPGNSVST	605
DB	610 KENLELNGSILLGGYGPAGSAPLQRFNSINSVAST	642
XX	ADP04701 standard; protein; 639 AA.	
AC	ADP04701;	
DT	29-JUL-2004 (first entry)	
DE	Sea squirt protein with tissue specific expression in development Seq296.	
KW	sea squirt; regeneration medicine; gene therapy; cell proliferation;	
KW	differentiation; reproduction; environmental measurement; water survey.	
OS	Ciona intestinalis.	
XX		
PN	JP2004057129-A.	
XX		
PD	26-FEB-2004.	
XX		
XX	31-JUL-2002; 2002JP-00222593.	
XX		



Qy	76	ELAGICAMGEPPVHVHQSQKHLHGLREELNAIVPYLEEMKKKKVVERWNOFVHVIE--QIK 134
Db	1120	DLAGEVLVQAOKEQVEVHKKKQEAQVOELQS---KCSGGERARAE--LNDPKVHKLQWNE 1175
Qy	135	KISSEIRPADVPFPKVPVDOSDLS--LRKLDLTKD-----LESIOKEKS---D 178
Db	1176	SVTCMLNEAEGKAIKLAKDVASLSSQLODTQELLOEBETROKLVNSTKRLQLEBERNSLOD 1235
Qy	179	RLKQVIEHLSNLSHSLCEVLGI-----DFKQTVYVHPSLDEAEG--SKNLSNNTT- 225
Db	1236	QLDEEMAEAKONLERHISTLNIQLSDSKKQLQDPASTV---EALIEGKKRFQKEIENLTQ 1291
Qy	226	-IERLAAANRLRMKIQRMKQLODPASSMLELWNLMDTPLEEQMF-----QNITCN 277
Db	1292	QYBEKAANYDKLEKTKRQLQELDDLVVDLONORQLVSNLEKKQKRFQDLIAEKNISSK 1351
Qy	278	IA----ASEQIEITPNTLSTDFLNVYSEVLRLEQL- KASKMKDLVLKKKAELEEHRRA 332
Db	1352	YADERDRAEAAREKETKALSARALBEALEAKEELERTNKM-----LKAEMEDLVSSK 1405
Qy	333	HLVCEGYAEFSEIATEAGAI DPSLVLEQIEBAHIATVKEAASRKDILEKVERWQVACE 392
Db	1406	DDVGKNVHELEKSKRALETQMEEMKTOLESEDELOQT-SDAKLRLUEV--NMQLKGGQFE 1462
Qy	393	EEAWLEDYNDKDDNRYNAGRAH---LTLKAEAKARTLV----NKIPGMVDVLRTKIAAWK 445
Db	1463	RDLQARDEQNEKRRQRLQRLQHEVETELEDERQORALAAAAKKKLEG--DLKDLLEQADS 1520
Qy	446	NERGKE-----DFYDGVSLSSMLDEYMFVVRQEBQE-KKQROROKKLQDQ 490
Db	1521	AIKGREBAIKQLRKLQAKMDQFORELEDAASRDEIFATAKENEKKAKSLEADLMQLOEJ 1580
Qy	491	LKAEQEAALYGS---KPSPSKPLSTKAPRMSGGANRRLSLGATWQ---PPKTDILHSK 544
Db	1581	LAARARAKQADLEKELEAELEASSLSCRNALQDEKRRLEARIAQLEEELEEEQGNMEAM 1640
Qy	545	SVRAAKKTEBIGTLPSSSSRGLDIAGLPIKKLFSNASTLTRETPPKPPFAQIITPGNSVSS 604
Db	1641	SDRVKATQAOAEQLS-----NELATERSTAQKNESARQ---QLERQNKELR 1683
Qy	605	TPVRPIPTNNTDENRTPKTFTTALNPKTPMTVTAPMCMAMTTPSLANKVSAVTPSVLYDK- 663
Db	1684	SKLHEMAGV---KSPFKSTIAALEAK-----IAQLEEQVEQEAEREQAATKSLKQDKK 1735
Qy	664	-PEVTLOEQDIDYGFERRLA 682
Db	1736	LKEIILLOVE-----DERKMA 1750

Qy	179	RLKQVIEHLNLSLSCEVLGI-----DFKQTVYEVHPSLDRAEG--SKNLSNTT-	225
Db	1351	QLDDEEAKQNLERHISTLNIQLSDSKKLQDFASTV----BALBEGKRFKOEIENLTQ	1406
Qy	226	-IERLAAANRLIEMKIORMKLODPASSMLNLWMLMDTPLBEOQMF-----QNI	277
Db	1407	QYBEKAAAYDKLEKTKRNLQQLDLDVLVDNQRLVSNLEKKORKFDOLLAEBKNISSK	1466
Qy	278	IA----ASQEITPEPNTLSTDFLNYVESEVLRLEQL-KASKMKDLVLKKKABLEEHRRA	332
Db	1467	YADRDRAEAEAREKETKALSARALEEALAEKEBELERTNM-----LKAEMEDLVSSK	1520
Qy	333	HLVGESEYAEBSIEAIEAGATDPSVLVLEQIEAHYATVKEEAPSRKDILEKVERWQACE	392
Db	1521	DDVGNVHLEKSKRALETQMEEMKTQBELEDELOAT-EDAKLRLEV--NMQALKGQFE	1577
Qy	393	EEAWELEDYNDKDNRYNAGRGAH--LTLKRAEKARTLV---NKTPGMVDLVRTKIAANK	445
Db	1578	RDLQARDEQNEEKRQLOQLQHEHYETELEDERKQRALAAAAKKLEG--DLXDLLEQADS	1635
Qy	446	NERGKE-----DFTYDGVLSUSMLDEYMFVVRQSEKQB-KKRQRDRQKKLQDO	490
Db	1636	AIKGRBEAIKQRLKLAQMKDFQRELEDARASRDEIFATAKENEKAKSLEADLMQLOBD	1695

QY 491 LKAEQALYGS---KPSPSKPLSTKKAPRHSNGMGNRRSLGATWQ---PPKTDILHSK 544  
 Db 1696 LAAERARKQADLEKEELAEELASSLSGRNALQDEKRLLEARIQLEEELEEEQGMMEAM 1755  
 QY 545 SVRAAKKTEEIGTILSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604  
 Db 1756 SDRVRKATQQAQLS-----NELATERSTAQKNESARQ---QLERQNKELR 1798  
 QY 605 TPRVPTNNTEDEENRTPKTFALTNPKTPTMTVTAPQMAMTSLANKVSATPVSIVYDK- 663  
 Db 1799 SKLHEMEGAV---KSKFKSTIAALEAK-----IALEEQVEQAEAREKQATKSLKQDKK 1850  
 QY 664 -PEVTLQEDIDYSFEERRLA 682  
 Db 1851 LKEILLQVE-----DERKWA 1865

RESULT 15

ABU10398  
 ID ABU10398 standard; protein; 1945 AA.  
 XX  
 AC ABU10398;  
 DT 20-AUG-2003 (first entry)  
 XX  
 DE Human smooth muscle myosin heavy chain (hSMMyHC) variant #1.  
 XX  
 KW Human; smooth muscle myosin heavy chain; hSMMyHC; inotropic;  
 KW antilasthmatic; hypertensive; hypotensive; uterine; antiinflammatory;  
 KW gynaecological; tocolytic; cardiac; vascular disorder;  
 KW pulmonary disorder; reproductive disorder; immunological disorder;  
 KW asthma; hypotension; hypertension; urinary incontinence;  
 KW irritable bowel syndrome; menstrual cramp; premature labor; glaucoma;  
 KW bronchioconstriction; cardiac malfunction.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003032018-A1.  
 XX  
 PD 13-FEB-2003.  
 XX  
 PF 10-AUG-2001; 2001US-00927597.  
 XX  
 PR 10-AUG-2001; 2001US-00927597.  
 XX  
 PA (CYTO-) CYTOKINETICS INC.  
 XX  
 XX Malik F, Beraud C, Freedman R, Craven A, Sakowicz R, Hartman J;  
 XX WPI; 2003-479585/45.  
 XX N-PSDB; ACA62125.  
 XX  
 XX New human smooth muscle myosin heavy chain (hSMMyHC) polypeptide and  
 PT nucleic acid, useful for diagnosing, preventing or treating e.g. asthma,  
 PT hypertension, incontinence, menstrual cramps, premature labor or cardiac  
 PT malfunction.  
 XX  
 PS Claim 10; Fig 2A-B; 80pp; English.  
 XX  
 CC The invention describes an isolated human smooth muscle myosin heavy  
 CC chain (hSMMyHC) polypeptide. The hSMMyHC protein, nucleic acid, or its  
 CC modulator, is useful for diagnosing, preventing or treating vascular,  
 CC pulmonary, reproductive or immunological disorders. In particular, the  
 CC hSMMyHC polypeptide, nucleic acid, or its modulator is useful for  
 CC diagnosing, preventing or treating e.g. asthma, hypotension,  
 CC hypertension, urinary incontinence, irritable bowel syndrome, menstrual  
 CC cramps, premature labor, glaucoma, bronchioconstriction, cardiac  
 CC malfunction, or other medical conditions related to hSMMyHC function. The  
 CC hSMMyHC polypeptide or nucleic acid is also useful for screening  
 CC therapeutic agents or hSMMyHC modulators, which may be used for treating  
 CC the above-mentioned diseases or disorders. This is the amino acid  
 CC sequence of a human smooth muscle myosin heavy chain (hSMMyHC) variant

SQ Sequence 1945 AA;  
 Query Match 6.9%; Score 240; DB 6; Length 1945;  
 Best Local Similarity 20.8%; Pred. No. 7.2e-08;  
 Matches 154; Conservative 145; Mismatches 313; Indels 128; Gaps 31;  
 QY 22 ELNIVWEGVGPDTTRDRMLLELEQ---ECLEVYRRKVDQANRSRAQL---RKAIAEGEA 75  
 Db 1182 ETVTLKALDETRSHQAQVQEMRQKHAQAEELTEQLEQFQKRAKANDKQKQLEKNA 1241  
 QY 76 ELAGICSAMGPPVHVHRSNOKLHGLREELNAIPVLEEMKKKKVERWQFVHVE-OIK 134  
 Db 1242 DLAGELRVLGAKQOEVEHKKKLEAQVQELQ---KCSGERARAE-LNDKVKHLQNEVE 1297  
 QY 135 KISSEIRPADFVPFVKVVDQSDLS--LRKLDELTKD-----LESLOKESK---D 178  
 Db 1298 SVTGMINEAEGKAIKLAKDVASLSQLQDTQELLQEETROKLVNVTKLQLEERNLSQD 1357  
 QY 179 RLKQVIEHLNLSLCEVIGI-----DFQOTVYEVHPSLDEAEG--SKNLSNTT- 225  
 Db 1358 QLDSEMEAKQNLERHISTLNQLSDSKKKLQDFASTV---EALBEGKRFPQKTEIENLTQ 1413  
 QY 226 -IERLAAANRLREMKIQRMQKLODPASSMELNMLMDTPLEEQOMF-----QNITCN 277  
 Db 1414 QYEEKAAAYDKLETKNRLQELDDLVVDLQNLQVLEKKQKRFQDLAEKNIISK 1473  
 QY 278 IA---ASEQITEPTNTLSTDFLNYVESEVLRLEQL-KASRMKDVLVKKKAELEHRRRA 332  
 Db 1474 YADERDRAEAREKETKALSARALEALEAKEELERTNKM-----LKAEMEDLVSSK 1527  
 QY 333 HLVGEEGVAEBSFEIAEAGIDPSLVLEQIEAHATVKEEAFSKDILEKVERWQNAE 392  
 Db 1528 DDVGKGVHEKSKRALETOMEEMKTOLEEELEDELOAT-EDAKLRLEV--NMQALKGOFE 1584  
 QY 393 BEAWLEDYKDDNRVYNACRGH---LTLKRAEKARTLV---NKTPGMVDVLTRTIAAWK 445  
 Db 1585 RDLQARDEQNEKRRQLQRLHEYTELEDERKQALAAAKKLEGG--DLKDLLEQDS 1642  
 QY 446 NERGKE-----DFTYDGVSLSSMLDEYMFVRQKEQE-KKRQDQKKLQDQ 490  
 Db 1643 AIKGRBBAIKQLKLAQMKDFORELEDAERASDEI FATAKENEKAKSLEADLMQLQSD 1702  
 QY 491 LKAEQALYGS---KPSPSKPLSTKKAPRHSNGMGNRRSLGATWQ---PPKTDILHSK 544  
 Db 1703 LAAERARKQADLEKEELAEELASSLSGRNALQDEKRLLEARIQLEEELEEEQGMMEAM 1762  
 QY 545 SVRAAKKTEEIGTILSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604  
 Db 1763 SDRVRKATQQAQLS-----NELATERSTAQKNESARQ---QLERQNKELR 1805  
 QY 605 TPRVPTNNTEDEENRTPKTFALTNPKTPTMTVTAPQMAMTSLANKVSATPVSIVYDK- 663  
 Db 1806 SKLHEMEGAV---KSKFKSTIAALEAK-----IALEEQVEQAEAREKQATKSLKQDKK 1857  
 QY 664 -PEVTLQEDIDYSFEERRLA 682  
 Db 1858 LKEILLQVE-----DERKWA 1872

Search completed: November 20, 2004, 00:39:17  
 Job time : 94 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:17:35 ; Search time 26 Seconds  
(without alignments)  
1759.978 Million cell updates/sec

Title: US-10-619-685-2  
Perfect score: 3496  
Sequence: 1 MSSAVKDLQHQSTTCDSLL.....DIDYFERRAIYILARQMV 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUS.COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	6.9	1857	4 US-09-917-254-91	Sequence 91, Appl
2	240	6.9	1972	4 US-09-538-092-1084	Sequence 1084, Ap
3	229	6.6	885	2 US-08-533-306A-4	Sequence 4, Appli
4	229	6.6	885	2 US-08-742-923A-4	Sequence 4, Appli
5	228.5	6.5	1972	4 US-08-875-435B-4	Sequence 4, Appli
6	221	6.3	1972	4 US-08-875-435B-3	Sequence 3, Appli
7	219.5	6.3	3210	4 US-09-538-092-1154	Sequence 1154, Ap
8	215.5	6.2	3248	1 US-08-353-700-1	Sequence 1, Appli
9	215.5	6.2	3248	5 PCT-US95-16216-1	Sequence 1, Appli
10	214.5	6.1	2482	1 US-08-328-254-6	Sequence 6, Appli
11	210	6.0	1939	4 US-09-538-092-915	Sequence 915, App
12	209	6.0	1937	4 US-09-538-092-918	Sequence 918, App
13	207	5.9	816	2 US-08-533-306A-6	Sequence 6, Appli
14	207	5.9	816	2 US-08-742-923A-6	Sequence 6, Appli
15	207	5.9	2101	1 US-08-466-390-4	Sequence 4, Appli
16	207	5.9	2101	1 US-08-470-950-4	Sequence 4, Appli
17	207	5.9	2101	1 US-08-467-781-4	Sequence 4, Appli
18	207	5.9	2101	1 US-08-195-487-4	Sequence 4, Appli
19	207	5.9	2101	2 US-08-483-924-4	Sequence 4, Appli
20	207	5.9	2101	3 US-09-452-294-1	Sequence 1, Appli
21	207	5.9	2101	5 PCT-US93-06160-4	Sequence 4, Appli
22	206.5	5.9	3878	4 US-09-914-259-11	Sequence 11, Appl
23	206	5.9	630	4 US-09-248-796A-20275	Sequence 20275, A
24	206	5.9	2349	4 US-09-538-092-914	Sequence 914, App
25	203.5	5.8	2662	4 US-09-595-684B-31	Sequence 31, Appl
26	203.5	5.8	2663	4 US-09-538-092-1252	Sequence 1252, Ap
27	199	5.7	976	4 US-09-538-092-1339	Sequence 1339, Ap

28	196.5	5.6	752	4 US-09-917-254-70	Sequence 70, Appl
29	194	5.5	1087	4 US-09-914-259-12	Sequence 12, Appl
30	194	5.5	1279	3 US-09-724-517-2	Sequence 2, Appli
31	194	5.5	1279	4 US-09-641-807A-2	Sequence 2, Appli
32	194	5.5	1279	4 US-09-723-096-2	Sequence 2, Appli
33	194	5.5	2954	4 US-09-150-867-1	Sequence 1, Appli
34	193	5.5	976	3 US-09-104-324B-4	Sequence 4, Appli
35	189	5.4	1093	5 PCT-US93-03077-1	Sequence 1, Appli
36	188.5	5.4	1960	4 US-09-538-092-1077	Sequence 1077, Ap
37	188	5.4	924	4 US-09-248-796A-18798	Sequence 18798, A
38	187.5	5.4	1976	4 US-09-538-092-1078	Sequence 1078, Ap
39	186.5	5.3	622	4 US-09-248-796A-19230	Sequence 19230, A
40	186.5	5.3	1078	4 US-09-248-796A-20284	Sequence 20284, A
41	185	5.3	525	4 US-09-107-532A-5095	Sequence 5095, Ap
42	184.5	5.3	1231	4 US-08-714-741-41	Sequence 41, Appl
43	184	5.3	1935	4 US-09-538-092-916	Sequence 916, App
44	183	5.2	1232	4 US-09-592-054-2	Sequence 2, Appli
45	183	5.2	1234	4 US-09-592-054-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-09-917-254-91  
; Sequence 91, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224(JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 91  
; LENGTH: 1857  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-917-254-91

Query Match		6.9%	Score 240;	DB 4;	Length 1857;
Best Local Similarity		20.8%	Pred. No. 1.2e-09;		
Matches 154;		Conservative 145;	Mismatches 313;	Indels 128;	Gaps 31;
QY	22	ELNVIWDEVGEPDTRDRMLLEQ---	ECLEVYRRKVDQANRSRAQL---	RKAIAEGEA	75
DB	1060	EVTVLKALDEETRSGHEAQVQEMRQKHAQAVEBELTEQLQFQKRAKANLDKNKQTLKENA	1119		
QY	76	ELAGTCSANGEPPVHVQRQSNOKLHGLREELNAIVPYLEEMKKKKVRRNQFVHVIE-OIK	134		
DB	1120	DLAGELRVLGQAQKEVHEHKKKLEAQVQELQS---	KCSGGERARAB-LNDKVHKLQNEVE	1175	
QY	135	KISSEIRPADFPFVKPVVDQSDLS--LRKLDLDTKO---	LESLOKEKS---	D 178	
DB	1176	SVTGMNLEAGKAIKAKDVASLSOQLOTOELLQETQKLVNVTKLQLEERNLSQD	1235		
QY	179	RLKQVIEHLNLSLCEVLGI-----	DFKQTVYEVHPSLDEAEG--	SKNLNTT--	225
DB	1236	QLDEEMEAQNLERHISTLNQLSDSKKLLQDFASTV---	EALBEGKKRFQKEIENLTQ	1291	
QY	226	-IERLAAAANRLREMKIQRMQKLODFASSMLELWNLMDTPLEEQMF-----	QNITCN	277	
DB	1292	OYEEKAAAYDKLEKTKNRLQQLDLDLVLDNRQVLNLEKQKQKQKPDQLLAEKKNISK	1351		
QY	278	IA-----ASROEITEPNTLTDFLNVYSESVLRLEQL--KASKMKDLVLKKAEELEHRRRA	332		
DB	1352	YADERDRABAEAREKETKALSARALEEALAEKEELERNKMK-----	LKAEMEDLVSSK	1405	
QY	333	HLVGBEGYAEFSEIAEAGAIIDPSLVLEQIEAHIAITVKEEAFSRKDIILEKVERMQNACE	392		

Db 1406 DDVGKVNHELEKSKRALETKQEMKTLQEELEDELQAT-EDAKLRLEV--NMQALKGQFE 1462  
QY 393 EEAWLEDYKDDNRYNAGRAH---LTLKRAEKARTLV---NKIPGMVDVLRTKIAAWK 445  
Db 1463 RDLOARDEQNEEKRRQLQRLHEYTELEDERKQALAAAKKLEG--DLKDLLELQADS 1520  
QY 446 NERGKE-----DFTYDGVSLSSMLDEYMFVRQKEQE-KKRQDQKQLQDQ 490  
Db 1521 AIKGREAIKQLKLOAQKMDQFQRELEDAARSDEIFATAKENEKKAKEADLMQLOED 1580  
QY 491 LKAEQALYGS---KPSKPLSTKAPRHSNGANRRLSLGGATMQ---PPKTDILHSH 544  
Db 1581 LAAERARKQADLEKEELABELASSLSGRNALQDEKRLLEARIQLEEELEEEQGNMEAM 1640  
QY 545 SVRAAKTTEIGTILSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604  
Db 1641 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1683  
QY 605 TPVRPITNTTDDENRTPKTFALTNPMTVTAPMQMAMTPSLANKVVSATPVSIVYDK- 663  
Db 1684 SKLHEMAGV---KSKFKSTIAALEAK-----IAQEEQVEQEAEREQAATKSLKQDKK 1735  
QY 664 -PEVTLOEDIDYSPERRLA 682  
Db 1736 LKEILLQVE-----DERKMA 1750

RESULT 2  
US-09-538-092-1084  
; Sequence 1084, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Glot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO 1084  
; LENGTH: 1972  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)-(0)  
; OTHER INFORMATION: Polypeptide Accession Number P35749  
US-09-538-092-1084

Query Match 6.9%; Score 240; DB 4; Length 1972;  
Best Local Similarity 20.8%; Pred. No. 1.3e-09;  
Matches 154; Conservative 145; Mismatches 313; Indels 128; Gaps 31;  
QY 22 ELNVTIWEVGEPTTTRDRMLLEQ---ECLEVYRKVDQNRSAQL---RKAIAEGEA 75  
Db 1175 EYTVLLKALDETRSGEAGVQEMRQKHAQVBELEQEFKRAKANLDKNTLEKENA 1234  
QY 76 ELAGICSAMGEPPVHVQRQSNKHLGLBELNAIVLYSEMKKKVERWQFVHVE-QIK 134  
Db 1235 DIAGELRVLGAKQAEVHHKKLELAQVQELQES---KCSGDERARAB-LNDKVHKLQNEV 1290  
QY 135 KISSIRPADFPVFPVPVQDSLS--LRKLDLTXD-----LESLOKEKS---D 178  
Db 1291 SVTGMLENEAGKAIKLANDVASLSQLQDTQELLQETFRQKLVNSTKLRLQEEERNSLOD 1350  
QY 179 RLKQVIEHLSLSICEVLGI-----DFKQTVVEVHPSLDEAEG--SKLSNVT- 225

Db 1351 QLDMEAEKQNLERHISTNLQLSDSKKKLQDPASTV-----EALLEGKKRFQKEIENLTQ 1406  
QY 226 -IERLAAAANRLREMKLQRMQKLODPASSMELMNLMDTPLEEQOMF-----ONITCN 277  
Db 1407 QYEEKAAAYDKLETKNRLQOELDDLVVDNQQLVSNLEKQKQKFDQLAEEKNISSK 1466  
QY 278 IA---ASEQIITPNTLTSTDPLNVESEVLRLSQL-KASKMKDVLKKALEBEHRRRA 332  
Db 1467 YADERDRAEAREKETKALSARALEAELEERTNKM-----LKAEMEDLVSSK 1520  
QY 333 HLVEEGYAEFEISIAEAGIDPSLVLEQIEAHIAIATVKEEAFSRKDIKLEKVERWQNAE 392  
Db 1521 DDVGKVNHELEKSKRALETKQEMKTLQEELEDELQAT-EDAKLRLEV--NMQALKGQFE 1577  
QY 393 EEAWLEDYKDDNRYNAGRAH---LTLKRAEKARTLV---NKIPGMVDVLRTKIAAWK 445  
Db 1578 RDLOARDEQNEEKRRQLQRLHEYTELEDERKQALAAAKKLEG--DLKDLLELQADS 1635  
QY 446 NERGKE-----DFTYDGVSLSSMLDEYMFVRQKEQE-KKRQDQKQLQDQ 490  
Db 1636 AIKGREAIKQLKLOAQKMDQFQRELEDAARSDEIFATAKENEKKAKEADLMQLOED 1695  
QY 491 LKAEQALYGS---KPSKPLSTKAPRHSNGANRRLSLGGATMQ---PPKTDILHSH 544  
Db 1696 LAAERARKQADLEKEELABELASSLSGRNALQDEKRLLEARIQLEEELEEEQGNMEAM 1755  
QY 545 SVRAAKTTEIGTILSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604  
Db 1756 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1798  
QY 605 TPVRPITNTTDDENRTPKTFALTNPMTVTAPMQMAMTPSLANKVVSATPVSIVYDK- 663  
Db 1799 SKLHEMAGV---KSKFKSTIAALEAK-----IAQEEQVEQEAEREQAATKSLKQDKK 1850  
QY 664 -PEVTLOEDIDYSPERRLA 682  
Db 1851 LKEILLQVE-----DERKMA 1865

RESULT 3  
US-08-533-306A-4  
; Sequence 4, Application US/08533306A  
; Patent No. 5837457  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Pu  
; APPLICANT: Collins, Francis S.  
; APPLICANT: Siciliano, Michael J.  
; TITLE OF INVENTION: Markers for Detection of Chromosome 16  
; TITLE OF INVENTION: Rearrangements  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
; STREET: P.O. Box 828  
; CITY: Bloomfield Hills  
; STATE: MI  
; COUNTRY: USA  
; ZIP: 48303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/533,306A  
; FILING DATE: September 25, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, DeAnn F.  
; REGISTRATION NUMBER: 36683  
; REFERENCE/DOCKET NUMBER: 2115-00869COB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 885 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-533-306A-4

Query Match 6.6%; Score 229; DB 2; Length 885;  
Best Local Similarity 21.2%; Pred. No. 2.9e-09;  
Matches 160; Conservative 140; Mismatches 311; Indels 142; Gaps 32;

```
QY 25 VIV-----DEVG--EPDTRDRMLLEQECLEVVRRKV-----DOANR-----SR 63
DB 107 VIVKGWIDQLRDGMGCLFDEBERAQOEDALAQAFEEARRRTREFDRSHREMEAK 166
QY 64 AOL---RKAIAEGEAEAGICSMGEPVHVHVRQSNQKHLGRLREELNAIVPYLEEMKKKV 120
DB 167 ANLDKNNQTLKENADLAGELRVLGQAQVEHKKKKLEAQVQELQS---KCSGGERARA 223
QY 121 ERNQVPHVIE--QIKKISSEIRPADVPKVPVVDQSDLS--LRKLDLTKD-----168
DB 224 E-LNDKVHKLQNEVESVTGMLNEAEGKAIKLAQVAVSLSSQLQDTQELLQOETROKLNVS 282
QY 169 --LESLOKES---DLKQVIEHLNLSHLSLCEVLGI-----DFKQTVVHVPSLD 213
DB 283 TKLRQLEERNLSQDQDEMEAKQNLERHISTLNQLSDSKKKLQDFASTV---EAL 338
QY 214 EABG--SKNLSNTT--IERLAAANRLREMKIQRMQKLDQFASMSLELNLMDTPLEEQ 269
DB 339 EGKRRFOKEIENLTQYEEKAAAYDKLETKNRLQOELDLVVDLNDQRLVSNLEKKQR 398
QY 270 MF-----QNTICNTIA---ASEQITEPTNTLSTDFLNVSEVLRLEQL-KASQKDL 317
DB 399 KFDQLLAEBKNISSKVADERDRAEAAREKETKALSARALEBALEKEELERTNKN---455
QY 318 VLKKALEBEHRRRAHLVGEEGYAEFFSIEAEGAIQPSLVLEQIEAHATVKEBAFSR 377
DB 456 ---LKAEMEDLVSSKDDVGVKNVHELEKSKRALETQMEEMKTQLELEDEL-QASEDAKLR 511
QY 378 KDILEKVERWQNAACEBEAMLEDYNDKDNRYNAGRGAH-----ITLKRAEKARTLVNKP 432
DB 512 LEV--NQALKQGFERDLQARDQNEEKRLQRLQHEYTELEDEDERNERALAAAKKL 569
QY 433 MVDVLRTKIAAKNRGKE-----DFTYDGVSLSSMLDEYMFVROEKEQE- 477
DB 570 EGDLDLELQADSAIKGREAIAKQLRKLQAKMDQFQLEDAASRDEIFATAKENEXKA 629
QY 478 KQRQDKKLDQDLKAEQEAALYGS---KPSKPLSTKKAPRHSMGANGRLSLGATMQ 534
DB 630 KSLADLMQLQEDLAAARARKQADLEKEBELASSUSGRNALQDEKRLREARTIAQLE 689
QY 535 ---PPKTDILHKSVAAKTEETIGTLPSSSRGLDIAGLPFKKLSFNASTLRETETPRK 591
DB 690 EBLEEGNMEASMDRVKATQAEQLS-----NELATERSTAQKNESARQ 735
QY 592 PPAQITPGNSVSTPVRPITNTEDDTPKTTFTALNPKTPMTVTAPMQMAMTPSLANK 651
DB 736 ---QLERQNKRLSKLHEMEGAV---KSKFKSTIAALEAK-----IAQLEEQVEQAREK 784
QY 652 VSATPSLVVDK--PEVTLQEDIDYSEERLA 682
DB 785 QAATKSLKQKDKKLKEILLQVE-----DERKMA 812
```

## RESULT 4

US-08-742-923A-4  
; Sequence 4, Application US/08742923A  
; Patent No. 5869611  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Pu  
; APPLICANT: Collins, Francis S.

APPLICANT: Siciliano, Michael J.  
APPLICANT: Claxton, David  
TITLE OF INVENTION: Markers for Detection of Chromosome 16  
TITLE OF INVENTION: Rearrangements  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: USA  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742,923A  
FILING DATE: No. 5869611 member 1, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Dean F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-00869DVC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-742-923A-4

Query Match 6.6%; Score 229; DB 2; Length 885;  
Best Local Similarity 21.2%; Pred. No. 2.9e-09;  
Matches 160; Conservative 140; Mismatches 311; Indels 142; Gaps 32;

```
QY 25 VIV-----DEVG--EPDTRDRMLLEQECLEVVRRKV-----DOANR-----SR 63
DB 107 VIVKGWIDQLRDGMGCLFDEBERAQOEDALAQAFEEARRRTREFDRSHREMEAK 166
QY 64 AOL---RKAIAEGEAEAGICSMGEPVHVHVRQSNQKHLGRLREELNAIVPYLEEMKKKV 120
DB 167 ANLDKNNQTLKENADLAGELRVLGQAQVEHKKKKLEAQVQELQS---KCSGGERARA 223
QY 121 ERNQVPHVIE--QIKKISSEIRPADVPKVPVVDQSDLS--LRKLDLTKD-----168
DB 224 E-LNDKVHKLQNEVESVTGMLNEAEGKAIKLAQVAVSLSSQLQDTQELLQOETROKLNVS 282
QY 169 --LESLOKES---DLKQVIEHLNLSHLSLCEVLGI-----DFKQTVVHVPSLD 213
DB 283 TKLRQLEERNLSQDQDEMEAKQNLERHISTLNQLSDSKKKLQDFASTV---EAL 338
QY 214 EABG--SKNLSNTT--IERLAAANRLREMKIQRMQKLDQFASMSLELNLMDTPLEEQ 269
DB 339 EGKRRFOKEIENLTQYEEKAAAYDKLETKNRLQOELDLVVDLNDQRLVSNLEKKQR 398
QY 270 MF-----QNTICNTIA---ASEQITEPTNTLSTDFLNVSEVLRLEQL-KASQKDL 317
DB 399 KFDQLLAEBKNISSKVADERDRAEAAREKETKALSARALEBALEKEELERTNKN---455
QY 318 VLKKALEBEHRRRAHLVGEEGYAEFFSIEAEGAIQPSLVLEQIEAHATVKEBAFSR 377
DB 456 ---LKAEMEDLVSSKDDVGVKNVHELEKSKRALETQMEEMKTQLELEDEL-QASEDAKLR 511
QY 378 KDILEKVERWQNAACEBEAMLEDYNDKDNRYNAGRGAH-----ITLKRAEKARTLVNKP 432
DB 512 LEV--NQALKQGFERDLQARDQNEEKRLQRLQHEYTELEDEDERNERALAAAKKL 569
QY 433 MVDVLRTKIAAKNRGKE-----DFTYDGVSLSSMLDEYMFVROEKEQE- 477
```

Db 570 EGDLDLEQADSAIKGEEAIKQLKLAQMKDFORELEADARASGRDEIFATAKENEKA 629  
QY 478 KKRQDKQLQDLKAEQALYGS---KPSPKPLSTKAPRHSWGCGANRRLSLGGATWQ 534  
Db 630 KSLADLMQLQDLAAAEARKQADLEKEELAEELASSLSGRNALQDEKRRLEARIQLE 689  
QY 535 ---PPKTDILHKSVAARKKTEIGTLPSSSRGDLIAGLPKLSFNASTRETETPRK 591  
Db 690 ESELEEQGNEAMSDRVKATQQAQOLS-----NEIATERSTAKNESARQ 735  
QY 592 PFAQITPGNSVSTPVRPITNNEDDENRTPKTFALNPKTPMTVTPAPQAMTSLANK 651  
Db 736 ---QLERQNKELRSKLHEMEGAV---KSKFKSTIAALEAK-----IAQLEEQVEQEAEREK 784  
QY 652 VSATPVSIVYDK--PEVTLOEDIDVSFEERRLA 682  
Db 785 QAATKSLKQDKKLEKLEILLQVE-----DERKWA 812  
  
RESULT 5  
US-08-875-435B-4  
; Sequence 4, Application US/08875435B  
; Patent No. 6593304  
; GENERAL INFORMATION:  
; APPLICANT: Hasegawa, Kazuhide  
; APPLICANT: Arakawa, Emi  
; APPLICANT: Oda, Shoji  
; APPLICANT: Matsuda, Yuzuru  
; APPLICANT: Takahashi, Katsuhito  
; APPLICANT: Sugahara, Michihiro  
; APPLICANT: Ishiyama, Haruo  
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING  
; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO  
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND  
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE  
; TITLE OF INVENTION: RECOMBINANT DNA  
; FILE REFERENCE: 07898-013001  
; CURRENT APPLICATION NUMBER: US/08/875,435B  
; PRIOR FILING DATE: 1997-07-25  
; PRIOR APPLICATION NUMBER: PCT/JP96/00134  
; PRIOR FILING DATE: 1996-01-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1972  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
US-08-875-435B-4

Query Match 6.5%; Score 228.5; DB 4; Length 1972;  
Best Local Similarity 19.6%; Pred. No. 1e-08;  
Matches 151; Conservative 159; Mismatches 316; Indels 143; Gaps 30;  
  
QY 4 AVKQDLHQ-MSTCDLSLL-----ELNVINWDEVEGPDTRDRMLLELEQ---ECLEYVR 53  
Db 1150 ALKTELEDLTDTATQQLRAKREQEVTVLKKALDEETRSHEAQVQEMRQKHTQVVELT 1209  
QY 54 KVDQANSRAQL---RKAIAEGEAEIAGICSAMEGPPVHYR----- 92  
Db 1210 EOLEQFKRAKANLDTKTQLEKENADLAGEFLVGLQAKQVEHKKKLEQLQELQSKCS 1269  
QY 93 -----OSNOKLHGLREBELNAIVPYLEEMKKKVKVERNQFVHVTEIQIKKTSSEIRPADF 145  
Db 1270 DGERARAEINDVKHKLQNEVESVTGMLSEAGKAIK-----LAKEVASIGSLQDQOE 1322  
QY 146 VPFKVPVQSDLSLRLKLBELTKDLESLOKESDRL--KQVIE-HUNSLHSICEVLGIDFK 202  
Db 1323 LIQETETROKLNVY-TKLRLQEDERNLSQELDEEAKQNLERHISTLN-----IQLSDSK 1377  
QY 203 QTVYEVHSLDEAGSKNLSNTTIERL-----AAAANRLREMKIQRMKQLQDPASSML 255  
Db 1378 KKLQDFASTVESLESGKKRFQKEIELSLTQQVEEKAAYDKLEKTKNRLQQLDLDLVLDL 1437

QY 256 ELWNLMDTFLPEEQWF-----QNTICNIA---ASOEITEPTNTLSTDFLNVYSESVL 304  
Db 1438 NQRLVSNLEKKQKFFDOLLAEEKNISSKYADEDRDAEAEAREKETALSARALEEAL 1497  
QY 305 RLEQL-KASKMKOLVLLKAELEHRRRAHLVBEQVAEFSIFAIEAGADPSSLVLEQI 363  
Db 1498 AKEELERTNKW-----LKAEMEDLVSSKDDVGNVHELEKSKRALETQMEEMKTLBEL 1551  
QY 364 BAHITATVKEEAFSRDKILEKVERWQACEEAEWLDEYNKDDNRYNAGRGAAH---LTLKRA 420  
Db 1552 EDELQAT-EDAKLLEV--NMQALKVQFERDQARDEQNEEKRRQLQQLHEYTELEDE 1608  
QY 421 EKARTLV---NKIPGMVDVLRTKIAAKNBERGKEDFTYDGVSVLSMMLDEYMFVROE--- 473  
Db 1609 RKQRALAAAKKLEG--DLKDLEQADSAIKGEEAIKQLLQQAQMKDQFQRELDARA 1666  
QY 474 -----KEQEKCR---QRDQKQLQDLKAEQALYGS---KPSPKPLSTKAPRHS 518  
Db 1667 SRDEIFATAKENEKAKSLEADLMQLQEDLAAAEARARKQADLEKEELAEELASSLSGRNA 1726  
QY 519 MGGANRRLSLGGATWQ---PPKTDILHKSVAARKKTEETIGTLPSSSRGDLIAGLPKK 575  
Db 1727 LQDEKRRLEARIQLEEEEBEEQNGMEASDRVKATQQAQOLS-----NE 1772  
QY 576 LSFNASTLRETETPRKPPAQITPGNSVSTPVRPITNNEDDENRTPKTFALNPKTPMT 635  
Db 1773 LATERSTAKNESARQ---QLERQNKELSKLQEMEGAV---KSKFKSTIAALEAK---- 1822  
QY 636 VTAPQAMTSLANKVSAFVSIVYDK--PEVTLOEDIDVSFEERRLA 682  
Db 1823 -IAQLEEQVEQEAEREKQAAAKALKQDKLKEMLLQVE-----DERKWA 1865

## RESULT 6

US-08-875-435B-3  
; Sequence 3, Application US/08875435B  
; Patent No. 6593304  
; GENERAL INFORMATION:  
; APPLICANT: Hasegawa, Kazuhide  
; APPLICANT: Arakawa, Emi  
; APPLICANT: Oda, Shoji  
; APPLICANT: Matsuda, Yuzuru  
; APPLICANT: Takahashi, Katsuhito  
; APPLICANT: Sugahara, Michihiro  
; APPLICANT: Ishiyama, Haruo  
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING  
; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO  
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND  
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE  
; TITLE OF INVENTION: RECOMBINANT DNA  
; FILE REFERENCE: 07898-013001  
; CURRENT APPLICATION NUMBER: US/08/875,435B  
; CURRENT FILING DATE: 1997-07-25  
; PRIOR APPLICATION NUMBER: PCT/JP96/00134  
; PRIOR FILING DATE: 1996-01-25  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1972  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-08-875-435B-3

Query Match 6.3%; Score 221; DB 4; Length 1972;  
Best Local Similarity 21.0%; Pred. No. 3.9e-08;  
Matches 159; Conservative 133; Mismatches 302; Indels 164; Gaps 32;  
  
QY 22 ELNVINWDEVEGPDTRDRMLLELEQ---ECLEYVRVRRVQANRSRAQLRKA---IAEGEA 75  
Db 1175 EVTVLKKALDEETRSHEAQVQEMRQKHTQVVELTQLEQFKRAKANLDSKQTLKENA 1234  
QY 76 ELAGICSAMEGPPVHVRSQNKU-----HGLUREELNAIVP 110

1235 DLAGELRVLGQAQKEVEHKKKLEVLQDLQSKSDGERARAEKSKVHLKNEVESVTG 1294  
111 YLEEMKKKKVVERWQVHVIEQIKKISSIRPADFPVPKVPVDSLDL-----SLRKLDEL 166  
1295 MLNEAGKAIK-----LAKDVASIGSOLQ-----DTQELQETROKLNVS 1336  
167 KLESQKESKES---DRLKQVIEHLSLSLCEVLGI-----DFKQTVVEVHPSLD 213  
1337 K-LRQLEDERNLQDQDDEMEAKQNLERHVSNTLNQLSDSKKQLQDFASTIEVME--- 1391  
214 EAGSKNLS-----NTTIERLAAANRLREMKIORMQKLODFASSMELNMLMDTPEE 267  
1392 --EGKKRLQKEMEGLSQVEEKAAYDKLEKTKNRLQQLDELVDLVDLDNORQLVSNLEKK 1449  
268 QQMF-----QNITCNIA-----ASEQIEPTNTLSTDFLNVYSEVLRLEQL-KASKMK 315  
1450 QKKFDQLAEENKISSYADERDRAEAREKETKALSARALEEAELEKELERTNM- 1508  
316 DLVLKKAELERHRRRAHLVGBEGYAEFSIEAIEAGADPSLVLRQIEAHIAIVKEEAF 375  
1509 -----LKAEMEDLVSSKDDVGKVVHELEKSKRALETQMEEMKTQLESEDDVQAT-EDAK 1562  
376 SRDKILEKVERWQNAEEAEWLDYNDKDDNRVYNAGRAH---LTLKRAEKARTLV---N 428  
1563 LRLEV--NQALKGQFERDLQARDEQNEEKRRQLQRLHEYTELEDERKQALAAAAKK 1620  
429 KIPGMVDVLRKTAANKNBERGKEDFTYDGVSLSSM-----LDEYMFVROE-----KEQ 476  
1621 KLEG--DLKDLQADSAIKGREBAIKQLRKLQAKMDFQRELDADARSDEIFATSKEN 1678  
477 EKRR-----ORDQKLOLQAEALYGS---KPSPKPLSTKKAPRHSNGANRRLSLGG 530  
1679 EKAKSLEADLMQLQSDLAARARKQADLEKEELAEELASSUGRNTLODEKRLERAI 1738  
531 ATMPPKPTDILHKSVRAAKTEBEIGTLPSSSRGLDIAGLPIKLSFNASTLRET---- 586  
1739 AQLE-----ELEEEQGNMEAMSDR-VRKATLQAEQLSNELATERSTAQKN 1783  
587 ETPRKPPAQITPGNSVSTPVRITNTNEDDENRPTKFTALNPKPTMTVTPMQAMTP 646  
1784 ESARQ---QLERQNKELRSKLQVEGAV---KAKLKSTVAALBAK-----IAQLEEQVEQ 1832  
647 SLANKVSATPVSLVYDK--PEVTLQEDIDYSPFERLA 682  
1833 EAREQAQATSKLKQKOKKLEKVLQVE-----DERKWA 1865

RESULT 7

US-09-538-092-1154  
; Sequence 1154, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO 1154  
; LENGTH: 3210  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P49454

US-09-538-092-1154

Query Match 6.3%; Score 219.5; DB 4; Length 3210;  
Best Local Similarity 20.6%; Pred. No. 1.1e-07;  
Matches 139; Conservative 124; Mismatches 256; Indels 155; Gaps 29;  
QY 1 MSSAVKQDLH-----QMTTCDLSLLLELNVIVDEVEPDTTRDRMLLELEOELEVEYRKV 56  
DB 2119 VSELLKDKTHLQEKLOSLKDSQALSITKCELENOIAQNLKEKELLVXSESLOA---RL 2175  
QY 57 DOANRSRAQLRAIAEGAEAGELAGISAMGEPVHVVRQSNOKLHGRLEELNATVPVLEENK 116  
DB 2176 SESDEYKUNVSKAL---EALV-----EKGEFALRLSSQTQEVVHQLRGIEKLRVLEAD 2228  
QY 117 KKKVERWQVHVIVIQIKK-----ISSEIRPADPV 146  
DB 2229 KQK-----LHIAEKLERENDSLKDKVENLERELQMSBENQELVILDAENSKAEVE 2281  
QY 147 PPKVPVDSLSLRKLDLTQDLSQKEKSDRLKQVIEHLSLSLSEVLGIDPKQTVY 206  
DB 2282 TLTKTQIEEMARSL-KVFEL--DLVTLRSEKENLTQIOEQKQQLSELDKLLS-SFKSLLE 2337  
QY 207 EVHPSLDEAE-GSKNLSNTTIERLAAANRLREMKIORMQKLODFASSMELNMLMDTFL 265  
DB 2338 E-----KEQAEIQIKESKTAVEML---QNQLKELN-EAVALCGDQEIEMKATEQSLDPPI 2389  
QY 266 BEQQMFQF-----ITCNIAASEQIEPTNTLSTDFLNVYSEVLRLEQLKASK-----MKD 316  
DB 2390 EEEHQLRNSIEKLRARLEADEKK-----QLCVLQQLKESSEHHDLLKG 2432  
QY 317 LVLKKAELERHRRRAHLVGBEGYAEFSIEAIEAGADPSLVLRQIEAHIAIVKEEAFS 376  
DB 2433 RVNLERELEIARTNOEHAALAEANSKGEVETLKAKIEGTQSLRGLELDVVVTIRSE--- 2489  
QY 377 KXDIILEKVERWQNAEEAEWLDYNDK--DNRVYNAGRAHLLTKRAEKARTLVNKPFGWVD 435  
DB 2490 KEDLTNELQEQERISE---LEIINSSPENILQEQEKEQKQVQMK---EKSTANEMIQTLK 2544  
QY 436 VLRTKIAAWKN-----RGKED-----FTYDGVSLSSMLDE---YMFVR----- 471  
DB 2545 ELNVERVALHNDQEAACKAEQNLSSQVECLELEKALQQLGLDEAKNNYIVLQSSVNGLIQ 2604  
QY 472 -----QEKEQKQRQDKLQDLKAEQALYGSKPSKPLSTKKAPRHSNGGA 522  
DB 2605 EVEDGKQKLEKDEEISRLKNQIQDQEQVLVSKLSQVEG-----EHQL-WK 2648  
QY 523 NRRLSLGGATMQ-PPKTDILHKSVRAAKTEBEIGTLPSSSRGLDIAGLPIKLSF--- 578  
DB 2649 EQNLERLNLTVELEQKIQVLQSKNASLQDTLTVLQSSYKNLENELELT--KMDKMSFVEK 2706  
QY 579 -NASTLRETETPRK 591  
DB 2707 VNMKTAKETELQRE 2720

RESULT 8

US-08-353-700-1  
; Sequence 1, Application US/08353700  
; Patent No. 5599919  
; GENERAL INFORMATION:  
; APPLICANT: YEN, TIMOTHY J.  
; APPLICANT: RATTNER, JEROME B.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A  
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,  
; TITLE OF INVENTION: AND METHODS OF USE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN  
; STREET: 1601 MARKET STREET, SUITE 720  
; CITY: PHILADELPHIA  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-2307



207 EVHPSLDEAE-GSKNLSNTTIERLAAANRLREMIKQIOMKQLODFASSMELNLMMDTLP 265  
2338 E-----KEQAEIQKEESKTAVEML-----QNQLKELN-EAVALCGQOEIMKATEQSLDPP 2389  
266 EQQMFQON-----ITCNIAASEQIEPTNTLSTDFLNYVESEVLRLEQKASK-----MKD 316  
2390 EBEHQRLNSIEKLRARLEADEKK-----QLCVLQQLKESEHHDLLKG 2432  
317 LVLKKAELHEHRRRAHLVGEGBAEFSEIAEAGADPSLVLEQIEAHIAIVTKKEAFS 376  
2433 RVENLERELEIARTQNEHAALAEANSKGEVETLKAKIEGMSQSLRGLEDVVTIRSE----- 2489  
377 RDIILEKVERWQACBEEAEWLVDYND-DNRYNAGRGHILTKRAEKARTLVNKPFGMVD 435  
2490 KENLTNELQKEQERISE---LEIINSFENILQKEQEKVQMK--EKSSTAMEMLQTLQK 2544  
436 VLRTKIAAWKNE-----RGKED-----FTYDGVSLSSMLDE-----YMFVR----- 471  
2545 ELNERVAALHNDQEAACKAEQNLSSQVECLELEKAKQLLQGLDEAKNNYIVLQSSVKGLIQ 2604  
472 -----QEKEQEKRRDQKQLODQKAEQALYGSKPSKPLSTKKAPRHSMGGA 522  
2605 EVEDGKQKLEKDEEISRLKNIQDQEQVLVSKLSQVEG-----EHQ-L-WK 2648  
523 NRRSLGGATMO-PPKTDILHSKSVRAAKKTEIEIGTLPSSSRGLDIAGLPIKLSF--- 578  
2649 EQNLERLNLTVELEQKIQLVQSKNASLQDTLEVLOQSSYKNLENELELT--KMDKMSFVEK 2706  
579 -NASTLRETETPRK 591  
2707 VNKMTAKETELQRE 2720

RESULT 10

US-08-328-254-6  
Sequence 6, Application US/08328254  
Patent No. 5710022  
GENERAL INFORMATION:  
APPLICANT: Zhu, Xueliang  
APPLICANT: Lee, Wen-Hwa  
TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,254  
FILING DATE: 24-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,239  
FILING DATE: 22-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cachryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-CJ 1191  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2482 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-328-254-6  
Query Match  
6.1%; Score 214.5; DB 1; Length 2482;  
Best Local Similarity 20.5%; Pred. No. 1.8e-07;  
Matches 138; Conservative 125; Mismatches 256; Indels 155; Gaps 29;  
1 MSSAVKDOQH---QMSSTCDLSLLELVNWDVEGPEPTDRMRMLLELEQECLEVYRRKV 56  
1391 VSELLKDKTHLQEKLSLEKDSQALSLTKCELENQIAQLNKEKELLVKSESLOA---RL 1447  
57 DOANRSRAQLKATAEAGEAELAGICSANGEPVHVHQSNQKHLGRLREELNAIVPLEEMK 116  
1448 SESDYKELNVSKAL---EALV---EKGEFALRSSTQSEVHQURRGIEKURVRIEAD 1500  
117 KKKVERWQFVHVIOIKK-----ISSEIRPADFV 146  
1501 KKQ-----LHIAEKLRERENDSLKDKVENLERELQMSSEENQELVILDENSQAEVE 1553  
147 PFKVPVDQSDLSRLKDLBTLKDLSELOKEKSDRLKQVIEHLSHLSLCEVLGIDFKQTVY 206  
1554 TLKTQIEEMARSL-KVPFL--DLVTLRSEKENLTQIOEKQQLSELDKLLS-SFKSLLE 1609  
207 EVHPSLDEAE-GSKNLSNTTIERLAAANRLREMIKQIOMKQLODFASSMELNLMMDTLP 265  
1610 E-----KEQAEIQKEESKTAVEML---QNQLKELN-EAVALCGQOEIMKATEQSLDPP 1661  
266 EQQMFQON-----ITCNIAASEQIEPTNTLSTDFLNYVESEVLRLEQKASK-----MKD 316  
1662 EBEHQRLNSIEKLRARLEADEKK-----QLCVLQQLKESEHHDLLKG 1704  
317 LVLKKAELHEHRRRAHLVGEGBAEFSEIAEAGADPSLVLEQIEAHIAIVTKKEAFS 376  
1705 RVENLERELEIARTQNEHAALAEANSKGEVETLKAKIEGMSQSLRGLEDVVTIRSE--- 1761  
377 RDIILEKVERWQACBEEAEWLVDYND-DNRYNAGRGHILTKRAEKARTLVNKPFGMVD 435  
1762 KENLTNELQKEQERISE---LEIINSFENILQKEQEKVQMK--EKSSTAMEMLQTLQK 1816  
436 VLRTKIAAWKNE-----RGKED-----FTYDGVSLSSMLDE-----YMFVR----- 471  
1817 ELNERVAALHNDQEAACKAEQNLSSQVECLELEKAKQLLQGLDEAKNNYIVLQSSVKGLIQ 1876  
472 -----QEKEQEKRRDQKQLODQKAEQALYGSKPSKPLSTKKAPRHSMGGA 522  
1877 EVEDGKQKLEKDEEISRLKNIQDQEQVLVSKLSQVEG-----EHQ-L-WK 1920  
523 NRRSLGGATMO-PPKTDILHSKSVRAAKKTEIEIGTLPSSSRGLDIAGLPIKLSF--- 578  
1921 EQNLERLNLTVELEQKIQLVQSKNASLQDTLEVLOQSSYKNLENELELT--KMDKMSFVEK 1978  
579 -NASTLRETETPRK 591  
1979 VNKMTAKETELQRE 1992

RESULT 11

US-09-538-092-915  
Sequence 915, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Giot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuraPatSeqFormatter Version 0.9

```
; SEQ ID NO 915
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P12882
US-09-538-092-915
```

```
Query Match          6.0%; Score 210; DB 4; Length 1939;
Best Local Similarity 20.9%; Pred. No. 2.8e-07;
Matches 138; Conservative 103; Mismatches 280; Indels 140; Gaps 23;
```

```
QY 3 SAVKQLHQMSTTCDSLLLELVNWIWVEGEPDTRDRMLLEQECLEVRKVDQANRS 62
Db 1331 SALAHALQSSRRHDCDLL-----REQYEEOQAKAELOQAMSKANSE 1371

QY 63 RAQLR-----KATAEGEAEIAGTCSAMGPPVHVROSQKHLGLR----- 102
Db 1372 VAQWRKTYETDAIQRTEELAEAKKLA--QRLQDAEEHVEAVNAKASLEKTKQRLQNE 1428

QY 103 -----BELNAIVPYLEEMK---KKVERWQNVFHVIE-QIKKISSEIRPADVPVKV 150
Db 1429 VEDLMIDVERTNAACAALDKKQRFDKTLAEWKQKCEETHAELEASQKESRSLSTELFKI 1488

QY 151 PVDQDLSLRKLDLTKDLESQKESDRKLOVIEHLNSLSLCEVLGIDPKQTVYEVHP 210
Db 1489 K-NAYEESLDQLETUKRENKNIQOISDLTEQIAGGKRIHEL-EKIKQVEQKSELQA 1546

QY 211 SDEAAGSKNLSNTTIERLAAANRLREMKIQRMOKLQDFASSMLELNLMDTPLEEQM 270
Db 1547 ALDEAEAS-----LEH-----EGKILRIQ-----LELNQV 1572

QY 271 FQNTICNTAAEQEITEPTNTLSTDFLNVYSEVLRLEQLKASKMKDLVLKKAELHEHRR 330
Db 1573 KSEVDRKIAEKDEEIDQ---MKRNHRIVESQSTLDAEIRSRNDALRLKKMEGDLNEM 1629

QY 331 RAHLVGEEGYAEFEFIEAIEAGADPSLVLEQIEAHATVKEEAFSRKDIILEKVERWONA 390
Db 1630 ETQL-----NHANRMAEALR-NYRNTQAILKDTQLHLDDALRSQEDLKEQLAMVERRANL 1684

QY 391 CEEEAWEEDYNKDDNRYNAGRAHUTLKRAEKARTLVNK-----IPGMVDVLRTKIAAWN 446
Db 1685 LQAE--IEE-----LRATLEQTERSRIKAEQELLIDASERVOLLHTQNTSLN 1729

QY 447 ERGKEDFTYDGVSLSSMLDEYMFVQEKEQKQRQDKLQDQKAEQALYGSKSPS 506
Db 1730 TKKLETDISIQI-Q-GEMEDIIQEARNABEKAKAITDAAMMAELKKEQDTSALHERMKK 1788

QY 507 KPLSTKKAAPRHSMGANRLSLGGATMPPKTDILHSHSVRAAKKTEIGTLPSSSRGL 566
Db 1789 NLEQTVKDLQHLRDEA-EQLALKGKKQIQKLE-----ARVRELEGEVESEQKENV 1838

QY 567 D-IAGL-----PIKKLSFNASTLRTEETPRKPPFAQITPGNSVSTPVPPINTTEDDNR 620
Db 1839 EAVKGLRKHERKVKELTY-----QTEEDRKNIILQLDVLQKQAKVSKYKQAEAEAEQ 1892

QY 621 T 621
Db 1893 S 1893
```

```
RESULT 12
US-09-538-092-918
; Sequence 918, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
```

```
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 918
; LENGTH: 1937
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P13535
US-09-538-092-918
```

```
Query Match          6.0%; Score 209; DB 4; Length 1937;
Best Local Similarity 20.1%; Pred. No. 3.3e-07;
Matches 134; Conservative 104; Mismatches 272; Indels 156; Gaps 21;
```

```
QY 4 AVKQDLHQMSTTCDSLLLELVNWIWVEGEPDTRDRMLLEQECLEVRKVDQANRSR 63
Db 1299 ALVSQLSRSKQAOSTQOIEELKHQLEETKAKNALAHALQSSRHDC-DLLREQYEEOEGK 1357

QY 64 AQLRKALAEAEALA-----GICSMGEPVHVHROSQKHLGL 101
Db 1358 AELQALSKANSEVAQWRKTYETDAIQRTEELAEAKKLAQRLQEAEEHVEAVNAKASL 1417

QY 102 R-----BELNAIVPYLEEMK---KKVERWQNVFHVIEQIKKISSEI 140
Db 1418 EKTQRLQNEVEDLMLDVERSNAAALDKKQRFDKVLSEWKQKYEETQAELEASQKES 1477

QY 141 RPADFVPKVPVDSDSLRLKLDLTKDLESQKESDRKLOVIEHLNSLSLCEVLGID 200
Db 1478 RSLSTELFKVK-NVYEESLDQLETUKRENKNIQOISDLTEQIAGGKQIHEL-EKIKQ 1535

QY 201 FKQTVYEVHPVPSLQAEAGSKNLSNTTIERLAAANRLREMKIQRMOKLQDFASSMLELNL 260
Db 1536 VEQEKCEIQAALEAEAS-----LEH-----EGKILRIQ----- 1565

QY 261 MDTPLEEQMFQNTICNTAAEQEITEPTNTLSTDFLNVYSEVLRLEQLKASKMKDLVLK 320
Db 1566 -----LELNQVSKSEVDRKIAEKDEEIDQ---LKRNTHRTVETMQSTLDAEIRSRNDALRVK 1618

QY 321 KKAELSHRRRAHLVGEEGYAEFEFIEAIEAGADPSLVLEQIEAHATVKEEAFSRKDI 380
Db 1619 KQEGDLNEMEIQ-----NHANRLAEALR-NYRNTQGIKLETQLHLDDALRGQEDLKEQ 1673

QY 381 LEKVERWQNAE---EBAWLEDYNKDDNRYNAGRAHUTLKRAEKARTLVNK-----IPGM 433
Db 1674 LAIVERRANLLQAEIEELW-----ATLEQTERSRIKAEQELLIDASER 1715

QY 434 VDVLRTKIAAWNKRGEKEDFTYDGVSLSSMLDEYMFVQEKEQKQRQDKLQDQK 492
Db 1716 VQLLHTQNTSLNT--KKKLENDVSQLQSEVEEVIQESRNAEAKAKAITDAAMMAELK 1773

QY 493 AEGEALYGSKSPSKPLSTKKAAPRHSMGANRLSLGGATMPPKTDILHSHSVRAAKKT 552
Db 1774 KEQDTSALHERMKKNLEQTVKDLQHLRDEA-EQLALKGKKQ----- 1814

QY 553 EEIGTLPSSSRGLDIAGLPKILSFNASTLR-ETETPRKPPFAQITPGNSVSTPVPPI 611
Db 1815 -----IQKLEARVRELEGEVENEQKRNAEAVKGLRKHERKVKELT 1854

QY 612 NNTEDD 617
Db 1855 YQTEED 1860
```

```
RESULT 13
US-08-533-306A-6
; Sequence 6, Application US/08533306A
```



Patent No. 5837457  
GENERAL INFORMATION:  
APPLICANT: Liu, Pu  
APPLICANT: Collins, Francis S.  
APPLICANT: Siciliano, Michael J.  
APPLICANT: Claxton, David  
TITLE OF INVENTION: Markers for Detection of Chromosome 16  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: USA  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/533,306A  
FILING DATE: September 25, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Deann F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-00869COB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 816 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-533-306A-6  
Query Match 5.9%; Score 207; DB 2; Length 816;  
Best Local Similarity 20.6%; Pred. No. 1.3e-07;  
Matches 153; Conservative 133; Mismatches 295; Indels 160; Gaps 30;  
QY 31 GEPDTRDRMLLEQECLEVVYR-----KVD-----QANRSRAQL 66  
DB 74 GEORQTPSREYVLDREAGVKYLKAPMLNGVCVIMQWIDQLQDGMGCLFDEERAQ 133  
QY 67 RKAIAGEAELAGICSAMGPPVHVRSQNKHLGLREEL-----NAIVPYLEEMKKKKVER 122  
DB 134 EDALQ-----QAFEBARRTRFEFDRDRSHREEMENEVESVTGMLNEAEGKAIK- 183  
QY 123 WNOFVHVIEQIKKISSEIRPAFVFPKVPVDSQL-----SLRKLDELTKDLESLOKEKS- 177  
DB 184 -----LAKDVASLSQLQ-----DTQELLQEETRQKLVNSTK-LRQLEBERNS 225  
QY 178 --DRLKQVIBHLSLHSLCBLVIGI-----DFKQTVYVHPSLDEAEG--SKNLSN 223  
DB 226 LQQLDDEMEAKONLHERHISTNLQISDSKKKLQDFASTV-----EALEEGKRFQKEIEN 281  
QY 224 TT--IERLAAAARLRMKIQKQKQDPASSMELMNLMDTPLEEQMF-----QNI 274  
DB 282 LTQOYEEKAAYDKLETKRNLQOELDLVVDLNQRLVSNLEKKQKQKFDQLLAEKNI 341  
QY 275 TCNIA-----ASEQIITEPNTLSTDFLNYVESEVLRLEQL-KASKMKDLVLKKXAELEHR 329  
DB 342 SSKYADERDRAEAAREKETKALSARALEEALEAKEELERTNKM-----LKAEMEDLV 395  
QY 330 RRAHLVGEGVAAEFSTEAETAGAIQSLVLEQIEAHVKEAFSRKDIKVERQW 389  
DB 396 SSKDDVGVKNVHELEKSKRALETOMEEMKLTQEELEDEL-QASEDAKLRLV--NMQALG 452  
QY 390 ACEEEAWLEDYNDKDDNRYNAGRGH-----LTLKRAEKARTLVNKPIMGVMDVLRTKIAAW 444

DB 453 QFERDLQARDEQNEKRRQLRQLHXYETELEDERNERALAAAAKKLEGDUKLELQAD 512  
QY 445 KNERGKE-----DFTYDGVSLSSMLDPEYMFVROEKEQ-KKQRDQKKLQD 489  
DB 513 SAIKGREEAIKQLRKLQAKMDQFQLEDEADARASRDEIFATAKENEKAKSLEADLMQLQE 572  
QY 490 QLKAEQEAALYGS---KPSPSKPLSTKAPRSHMGANRRSLSGGATMQ---PPKTDILHS 543  
DB 573 DLAAAARAKQADLEKEELABELASSLSGRNALQDEKRLLEARIAQLEEELEEEQGNMEA 632  
QY 544 KSVRAAKTEETIGTILSPSSSRGLDIAGLPIKKLSFNASTLTRETETPRKPPAQTITPNSVS 603  
DB 633 MSDRVKATQQAQSL-----NELATERSTAQKNESARQ---QLERQNKEL 675  
QY 604 STPVRPITNNTEDDENRTPKTFTALNPKTPMTVTAPMQMAMTPSLANKVYSATPVSIVYDK 663  
DB 676 RSKLHEMGA---KSKFKSTIAALEAK-----IAQLEEQVEQEAEREQAATKSLKQK 727  
QY 664 --PEVTLOQDIDYSFEERRLA 682  
DB 728 KLKEILLQVE-----DERKMA 743  
RESULT 14  
US-08-742-923A-6  
Sequence 6, Application US/08742923A  
Patent No. 5869611  
GENERAL INFORMATION:  
APPLICANT: Liu, Pu  
APPLICANT: Collins, Francis S.  
APPLICANT: Siciliano, Michael J.  
APPLICANT: Claxton, David  
TITLE OF INVENTION: Markers for Detection of Chromosome 16  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: USA  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742,923A  
FILING DATE: No. 5869611ember 1, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Deann F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-00869DVC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 816 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-742-923A-6  
Query Match 5.9%; Score 207; DB 2; Length 816;  
Best Local Similarity 20.6%; Pred. No. 1.3e-07;  
Matches 153; Conservative 133; Mismatches 295; Indels 160; Gaps 30;  
QY 31 GEPDTRDRMLLEQECLEVVYR-----KVD-----QANRSRAQL 66

```

Db 74 GEORQTPSYVDLREAGKYLKAPMILNGVCVIWGLDQLRDMGCLFDFBERAQ 133
QY 67 RKAIAEGEAEAGISAMGEPVHVHVSQKLGHEEL-----NAIVPVEEMKKKVER 122
Db 134 ERLAQ-----QAFEAARRRREFDRDRSHREEMENEVSTGMLNEAGKAIK- 183
QY 123 WQFVHVBIQIKKISSIRPADFPVFPKVPVDSLDL-----SLRKLDELTKDLESLOKES- 177
Db 184 -----LAKDVASLSQLQ-----DTQELLQEBTRQKLVNYSK-LFQLEBEENS 225
QY 178 --DRLKQVTEHLNSLHSLCEVLGI-----DFKQTVYEVHPSLDEAG--SKNLSN 223
Db 226 LQDQLEENEAQNLERHISTNLQISDSKKLQDFASTV-----BALBEGKKRFQKEIEN 281
QY 224 TT--TERLAAAANRLREMIORMOKLQDFASSMELMNLMDTPLEEQMF-----QNI 274
Db 282 LTQVVEEKAANDKLETKNRLQELLDLVLDNQRLVSNLEKKQKRFQDLAEEKNI 341
QY 275 TCNIA-----ASQEITEPNTLSTDFLNVYSEVLRLEQ--KASKMKDLVLKKAELBEHR 329
Db 342 SSKYADERDRAEAREKETKALSARALEAEAELEERTNKM-----LKAEMEDLV 395
QY 330 RAHLVGECEGYAEBSIESIAEAGADPSLVLEQIEAHIAITVKEBAFSPKDIILEKVERQW 389
Db 396 SSKDDVGKNVHLEKSKRALETQMEEMKTQLEBEDEL-QASEDAKLRLV--NMQALKG 452
QY 390 ACEEEAWLEDYKNDNRNAGRGH-----LTLKRAEKARTLVNKPIMVVDVLRTKIAAW 444
Db 453 QFERDQARDEONEKRRQLQOLHEYTELEDERNERALAAAANKKLEGDLKLELQAD 512
QY 445 KNERGKE-----DFTYDGVSLSSMLDEYMFVROKQE--KXQRQDKLQD 489
Db 513 SAIKGRERAIKQLKLAQMKDFORELEDAASRDEIFATAKENEKKALEADLMQLQE 572
QY 490 QLAKEQEAALYGS--KPSPSKPLSTKAPRSHMGANRLSLGGATMQ---PPKTDILHS 543
Db 573 DLAAERARKQADLEKEELAELASLSGRNALQDEKRLERLARIAQLEEELEEEQGNMEA 632
QY 544 KSVRAAKTEETIGTILSPSSSRGLDIAGLPKIKLSFNASTLRETETPRKFPFAQITPGNSVS 603
Db 633 MSDRVKATQAEQLS-----NELATERSTAQNESARQ---QLERQNKEL 675
QY 604 STVPPIPTNNTBDDNRTPKTFALNPKTMTVTAPMQMAMTPSLANKVSATPVSIVYD 663
Db 676 RSKLHEMGAIV--KSKFKPTIAALEAK-----IAQLEEQVEQEAEREQAATKSLKQDK 727
QY 664 --PEVTIQEDIDYSEERRLA 682
Db 728 LKLEILLQVE-----DERQMA 743

```

```

RESULT 15
US-08-466-390-4
; Sequence 4, Application US/08466390
; Patent No. 5686562
; GENERAL INFORMATION:
; APPLICANT: TOUNKATLY, GARY
; APPLICANT: TOUNKATLY, GARY
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,390
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ. EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-390-4

```

```

Query Match 5.9%; Score 207; DB 1; Length 2101;
Best Local Similarity 21.1%; Pred. No. 5.4e-07;
Matches 123; Conservative 95; Mismatches 137; Indels 168; Gaps 24;

QY 6 KQQLHQMSTTCDSLLLEINLVNDEVGEPDTRDRML-----LEBQECLEVYRRKVDQANR 61
Db 461 KQQLSSLTITDQSSISNLSQAKEBEQASQAHGARLTAQVASTSE-LTTLNATIQDQDQ 519
QY 62 SRAQLRKAIAEGEAEAGISAMGEPVHVHVSQKLGHEELNAIVPVEEMKKKVE 121
Db 520 ELAGLQKQAEKQALQAT-----LQQEQASQGLRHQVEQL---SSSLKQKE-- 564
QY 122 RWNQFVHVEIQIKKISSE---IRPADFPVFPKVPVDSLDLRLKDELTKDLESLOKESD 178
Db 565 -----QQLKEVAEQEATRODHAQQAQATAAEEREASLRERDAALKQLEALEKAA 615
QY 179 RLKQVIEHLNSLHSLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNTTIERLAAANRLRE 238
Db 616 KLEITLQQLQ-----VANEARDSAQTSTVTAQREKAELSR--- 650
QY 239 MKTQRMQKLDQFASSMELMNLMDTPLEEQMFQ-----NITCNIAASEQITEPNTLSTD 294
Db 651 -KVEELQA-----CVETARQEQHEAQVAEQLQLRSEQQKATEKERV AQ- 695
QY 295 FLNVYSEVLRLEQLKASKMKDLVLKKAELERHRRRAHLVGBEGYAEBSIESIAEAGAI 354
Db 696 -----EKDQLQ-EQLQA--LKESLKVTKGSLSEKRA----- 725
QY 355 DPSLVLEQIEAHIAITVKEBAFSPKDIILEKVERQWQNAACEEAWLEDYKNDNRNAGR--- 411
Db 726 --ADALEEQQCISELKAET---RSLVEQHKKRERKELEEE-----RAGRKGL 767
QY 412 -GAHLTLKRAEKARTLV-----NKIPGMVDVLRTKIAAKNKGKEDFTYDGV 458
Db 768 EARLLQLGEAHQAEFVLRRELAEAMAAQHTASECEQLVKEVAAMRD--GYDEMQQEEA 825
QY 459 SLSSMLDEYMFVROKQEKKRQDOKLQDLKAEQEAALYGSKPSKPLS---TKKAP 515
Db 826 QYGMFQEQMLTLKE-ECEKARQ-----ELQEAKEKVAGIESHSELQISROQNKLA 876
QY 516 RHMGGANRLRLSLGGATMQPPKTDILHKSVSRAAKTEETIGTL 558
Db 877 LHA-----NLARALQ-----VOEKEVRAQKLADLSTL 905

```

Search completed: November 20, 2004, 00:43:11  
Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:37:56 ; Search time 85 Seconds

(without alignments)  
2874.683 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKQDLHQMSTTCDLSL.....DIDSFERRLAIFYLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3496	100.0	690	15	US-10-619-685-2
2	2975.5	85.1	700	15	Sequence 2, Appl1
3	2939	84.1	690	17	Sequence 62819, A
4	1314.5	37.6	582	15	Sequence 253246
5	1265.5	36.2	576	16	Sequence 268883
6	1260	36.0	582	17	Sequence 131367
7	1233.5	35.3	511	15	Sequence 223724
8	1225	35.0	602	15	Sequence 49169, A
9	1213	34.7	630	16	Sequence 254332
10	1182.5	34.0	598	17	Sequence 200324
11	1182.5	33.8	589	16	Sequence 194860
12	1181.5	33.8	559	16	Sequence 135276
13	1141	32.6	660	17	Sequence 179742
					Sequence 287895

14	1127.5	32.3	610	16	US-10-437-963-106958
15	1101.5	31.5	611	17	US-10-425-115-264132
16	1079	30.9	578	17	US-10-425-115-354402
17	1077.5	30.8	953	17	US-10-425-115-284613
18	1074	30.7	603	15	US-10-425-114-64857
19	1060.5	30.3	598	17	US-10-425-115-264129
20	1042	29.8	553	16	US-10-437-963-147340
21	960.5	27.5	1010	17	US-10-425-115-312927
22	928	26.0	709	16	US-10-437-963-128829
23	907.5	26.0	185	16	US-10-437-963-186153
24	895.5	25.6	606	15	US-10-437-963-186153
25	874.5	25.0	554	15	US-10-425-114-65219
26	873	25.0	176	16	US-10-437-963-134843
27	870	24.9	538	16	US-10-437-963-175753
28	799	22.9	234	15	US-10-424-599-281618
29	709.5	20.3	147	16	US-10-437-963-183850
30	679.5	19.4	354	15	US-10-425-114-42188
31	671.5	19.2	220	17	US-10-425-115-287893
32	632	18.1	290	15	US-10-424-599-268884
33	546.5	15.6	569	17	US-10-425-115-259791
34	528.5	15.1	180	15	US-10-424-599-186497
35	489	14.0	315	16	US-10-767-701-42467
36	479.5	13.7	251	16	US-10-767-701-56338
37	478	13.7	195	17	US-10-425-115-195661
38	470.5	13.5	141	17	US-10-425-115-216368
39	457	13.1	125	15	US-10-424-599-226892
40	430.5	12.3	325	16	US-10-767-701-41154
41	429.5	12.3	164	16	US-10-767-701-50583
42	394	11.3	143	17	US-10-425-115-257687
43	382.5	10.9	222	15	US-10-424-599-148583
44	350	10.0	172	15	US-10-424-599-261176
45	323.5	9.3	620	17	US-10-733-878-431

ALIGNMENTS

RESULT 1  
US-10-619-685-2  
; Sequence 2, Application US/10619685  
; Publication No. US20040096875A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Resources, Ministry of Agriculture,  
; APPLICANT: Forestry and Fisheries  
; APPLICANT: Japan Science and Technology Corporation  
; TITLE OF INVENTION: A NOVEL GENE FOR CONTROLLING LEAF SHAPES  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/619,685  
; CURRENT FILING DATE: 2003-07-15  
; PRIOR APPLICATION NUMBER: US/09/667,475D  
; PRIOR FILING DATE: 2002-11-13  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-619-685-2

Query Match	100.0%;	Score	3496;	DB	15;	Length	690;
Best Local Similarity	100.0%;	Pred. No.	7.1e-216;	Mismatches	0;	Indels	0;
Matches	690;	Conservative	0;				
QY	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1
DB	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1
QY	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61
DB	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQNLHGLREELNAIVYLEMKKKV	61
QY	121	ERNWQFVHTEQTKTSSEIRPADFPVFKVPDQSDLSLRKLBELTKDLSLQEKSDRL	121	ERNWQFVHTEQTKTSSEIRPADFPVFKVPDQSDLSLRKLBELTKDLSLQEKSDRL	121	ERNWQFVHTEQTKTSSEIRPADFPVFKVPDQSDLSLRKLBELTKDLSLQEKSDRL	121

```

Db 121 ERWNOFVHVIEQIKKISSEIRPADFPVKVPVDSLSLRKLDLTKDLESQKESDRL 180
Qy 181 KOVIEHLNSLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Db 181 KOVIEHLNSLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Qy 241 IQRMQKLDQFASSMELNWLMDTPLEEQMFQNTICNIAASEQITEPTNTLSTDFLNVE 300
Db 241 IQRMQKLDQFASSMELNWLMDTPLEEQMFQNTICNIAASEQITEPTNTLSTDFLNVE 300
Qy 301 SEVLRLQKASQKMDLVKKAELEHRRRAHLVGEERYAEFSEIAEAGADPSLVL 360
Db 301 SEVLRLQKASQKMDLVKKAELEHRRRAHLVGEERYAEFSEIAEAGADPSLVL 360
Qy 361 EQIEAHATVKEEAFSRKDIKVERWQNAACEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Db 361 EQIEAHATVKEEAFSRKDIKVERWQNAACEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Qy 421 EKARTLVNKPICMVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVQEKEQKRR 480
Db 421 EKARTLVNKPICMVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVQEKEQKRR 480
Qy 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Qy 541 LHSKSVRAAKTEETIGTLPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db 541 LHSKSVRAAKTEETIGTLPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Qy 601 SVSSTPVRPITNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 660
Db 601 SVSSTPVRPITNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 660
Qy 661 YDKPEVTLQEDIDYSEFERRLAIVLARQMV 690
Db 661 YDKPEVTLQEDIDYSEFERRLAIVLARQMV 690

```

RESULT 2

```

US-10-425-114-62819
; Sequence 62819, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62819
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3332-041-P11_FLI pep
US-10-425-114-62819

```

```

Query Match 85.1%; Score 2975.5; DB 15; Length 700;
Best Local Similarity 85.1%; Pred. No. 1.9e-182;
Matches 588; Conservative 47; Mismatches 51; Indels 5; Gaps 5;
Qy 1 MSSAVKQDLQHSMTTCDLSLLELNVWDEVGEPDTRDRMLLELLEQEGLEVYRRKVDQAN 60
Db 12 MSSAVKQDLQHSMTTCDLSLLELNVWDEVGEPDTRDRMLLELLEQEGLEVYRRKVDQAN 71
Qy 61 RSRAQLRKAIRAGEAELAGICSAMGEPFVHVRSQNKQLHGLREELNAIVPYLEEMKKKV 120

```

```

Db 72 RCRAQLRKAIRAGEAELAGICSAMGEPFVHVRSQNKQLHGLREELNAIVPYLEEMKKKV 131
Qy 121 ERWNOFVHVIEQIKKISSEIRPADFPVKVPVDSLSLRKLDLTKDLESQKESDRL 180
Db 132 ERWNOFVHVIEQIKKISSEIRPADFPVKVPVDSLSLRKLDLTKDLESQKESDRL 191
Qy 181 KOVIEHLNSLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Db 192 KOVIEHLNSLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 251
Qy 241 IQRMQKLDQFASSMELNWLMDTPLEEQMFQNTICNIAASEQITEPTNTLSTDFLNVE 300
Db 252 VORMQKLDQFASSMELNWLMDTPLEEQMFQNTICNIAASEQITEPTNTLSTDFLNVE 311
Qy 301 SEVLRLQKASQKMDLVKKAELEHRRRAHLVGEERYAEFSEIAEAGADPSLVL 360
Db 312 SEVLRLQKASQKMDLVKKAELEHRRRAHLVGEERYAEFSEIAEAGADPSLVL 371
Qy 361 EQIEAHATVKEEAFSRKDIKVERWQNAACEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Db 372 EQIEAHATVKEEAFSRKDIKVERWQNAACEEAWLEDYKNDNRYNAGRAHLTLKRA 431
Qy 421 EKARTLVNKPICMVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVQEKEQKRR 480
Db 432 EKARTLVNKPICMVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVQEKEQKRR 490
Qy 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db 491 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 550
Qy 541 LHSKSVRAAKTEETIGTLPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db 551 LHSKSVRAAKTEETIGTLPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 609
Qy 601 SVSSTPVRPITNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 659
Db 610 NVSPMTPRPISSATE-BENKTPRTFVGLNAKTPATVTAPMQMAMTPAVANKVIATPATL 668
Qy 660 YDKPE-VTLQEDIDYSEFERRLAIVLARQMV 689
Db 669 FQEKADSPALPADIEYSFEERRLAIVLARQV 699

```

RESULT 3

```

US-10-425-115-253246
; Sequence 253246, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253246
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(690)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162540C.1.pep
US-10-425-115-253246

```

```

Query Match 84.1%; Score 2939; DB 17; Length 690;
Best Local Similarity 84.4%; Pred. No. 4.1e-180;

```

Matches	584;	Conservative	49;	Mismatches	53;	Indels	6;	Gaps	6;
Qy	1	MSSAVKQDLQHWSTTCDSLLLELVN	WDEVGEPTTDRRMLELELE	EQECLEVYRRKVDQAN	60				
Db	1	MSSAAKQDLQWSTTYSLLLELVN	WDEVGEPMARDMLELELE	EQEWLDVYRRKVDQAN	60				
Qy	61	RSRAQLKAIAGEAEALAGISAMGEP	PHVYRQSKLHGIREELNAV	PVYLEEMKKKV	120				
Db	61	RCRAQLKQAIADAEEALAGISAMSEP	PIHVYRQSKLHGIREELNAV	PVYLEEMKKKV	120				
Qy	121	ERNWQFVHVIBQIKKISSEIRPAD	PVPPKVP	-VDQSDLSLRKLDLT	KDLESLOKEKSDR	179			
Db	121	ERNWQFVDVIBQIKKVASEIRP	SDFVPII	IXVDQSDLSLRKLDLT	KELQSLQKEKSDR	180			
Qy	180	LKQVIEHLNLSHSLCEVLGIDFKQ	TVYVHPSPDSEAGSKNS	NTTIERLAAAANRLREM	239				
Db	181	LKQVMEHLSTLHSLCEVLGVD	FKQTVNEVHPSLGEADG	SKNSNCTISLSAASRLREL	240				
Qy	240	KIORMQKLODFASSMLELWNLMD	TPLESQOMFQNTC	NIASEQEIITEPNTLSTDFLNVV	299				
Db	241	KVQRMQKLODLASSMLELWNLMD	TPLESQOMFQNTC	NIASEHEIITEPNTLSTDFLSYV	300				
Qy	300	ESEVLRLBQLKASKMKDLVLKKA	ELEHRRRAHLVGEEGYAE	EESIEAIEAGADPSLV	359				
Db	301	ESEVLRLBQLKASKMKDLVLKKT	ELEHRRRAHLVGEEGYAE	ESDEAIEAGADPSLV	360				
Qy	360	LEQIEAHATLVKEEAFSRKDI	ILEKVERWQNACEEAEWL	DYNDKDNRYNAGRAHLLTKR	419				
Db	361	LEQIEAHATLVKEEAFSRKDI	ILEKVERWLNACEEEAEWL	DYNDKDNRYNAGRAHLLTKR	420				
Qy	420	AEKARTLVNKIPGMVDVLRTK	IAAKNBRGKEDFTYD	GVSLSSMLDEYWFVRYEQEKK	479				
Db	421	AEKARILVNKIPGLVDVLT	TKIVAEABERGKE	-FTYDGVRLMSMLEEYMIVRQEKLOKK	479				
Qy	480	RORDOKKLODQLKAEQEALY	GSKPSPKPLSTKKA	PRHSMGGANRLSIGGATMQPKTD	539				
Db	480	RORDOKKITODQLKAEQEALY	GSKPSPKPSQSTKKA	PRHSMGGANRLSIGGATMQAPKTD	539				
Qy	540	ILHSKSVAAKKTTEIGTILSP	SSSGGLDLAGLPIKKLS	FNASTLTRETETPRKFPQAITPG	599				
Db	540	ILHSKTAARAAKADGLGAL	S-SSRGLDLAGLPIKKLS	FNASTLTREAEETPRKFPQAIMP	598				
Qy	600	NSVSTPVRPIPTNTDEDDENR	TPKFTT	-ALNPKTPMTVTPMQAMT	PSLANKVSATPVS	658			
Db	599	NNVSPMPTRPISSATE	-EENKPTRFVUGLNAKT	PTATVTPMQAMT	PAVKVATPAT	657			
Qy	659	LVYDKPE-VTLQEDIDYSP	EEERLLAIYLARQM	689					
Db	658	LFOEKADSPALPADIEYSP	EEERLLAVYLARQV	689					

```

RESULT 4
US-10-424-599-268883
; Sequence 268883, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 268883
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84823C.1.pep
US-10-424-599-268883

```

Query Match	37.6%	Score	1314.5	DB	15	Length	582		
Best.Local Similarity	49.6%	Pred. No.	6.2e-76						
Matches	275	Conservative	100	Mismatches	140	Indels	39	Gaps	9

  

Qy	15	TCDSLLLELVNIWDEVEGPTTTRDRMLLELEQECLEVYRRKVDQANRSRAQLRAKIAEGE	74
Db	15	TCGSLKKLQEIWDEGESDEQRDKMLLQLEQCLDYYKRVQEAQKSRALQLQASDAK	74
Qy	75	AEIAGICSAMGEPVHVQRNQKLGHURBELNALVPYLEMKKKVWRNQFVHVIEQIK	134
Db	75	LELSTLLSALGKSFAGIPENTS-GTIIKEQLAATAPVLEQLWQOEKRIKFSDFVQSQIQ	133
Qy	135	KISSEI-----RPADEVFPKVPYDQSDLSLRKLDLTKDLESLOKESKDRLLKQVIE	185
Db	134	QICEIAGNLLNDVSPA-----VDESLSLKKLDEYQSELEQLOKESERLHKVLE	185
Qy	186	HLMSLSLCEVLGIDFKQTVYVHPVLSDEAEG--SKNLSNTTIERLAAAANRLREMKTOR	243
Db	186	FVSTVHDLCAVLGMDFTTATEVHPVLSNDSTGVQSKSISNTDLARLAKTVLTLEDKQKOR	245
Qy	244	MQKLQDPASSMLELWNLMDTPLERQQQFONITCNIAASEQIETPTNTLSTDFLNVYSEV	303
Db	246	LHKLOELASQLIDLWNLMDTHPERRFLFDVHTCNMSASVDVTPGALALDIEQAEVEV	305
Qy	304	LRLEQLKASKMKDLAVLKKKAELEBHRRAHLVGBEGVAEBSFISIAEAGADPSLVLEQI	363
Db	306	ERLDQLKASKRKEITAFKKQAELEIIFARAHIETVPDAAREKIMALIDSGNTEPTELLADM	365
Qy	364	EAHIATVKEAFSRKQILEKVERWONACEEBAWLEDYKNDNRYNAGRGAHULTKRAEKA	423
Db	366	DNQITAKAKEEALS RKDILDKVKWMSACEEBSWLEDYNRDRNRYNAGRIHLKRAEKA	425
Qy	424	RTLVNKIPGMVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKQRD	483
Db	426	RILVNKIPALVDITLVAKTRAWEEHDHSM-SFTYDGVPLMLDDEYAMLNRHEREEKKRMRD	484
Qy	484	QKKLDOLKAPQEAALYGSKPSKPLSTKKA--PRHSMGGA---NRRLSLGGATMQPPK	537
Db	485	QKKHHEQRNTEQETIFGSRPSPAPVSSSKGGPR-ANGGANATPNRRLSLNA-----	536
Qy	538	TDILHKSIVRAAKK	551
Db	537	-----HONGNRSTSK	546

RESULT 5  
US-10-437-963-131367  
; Sequence 131367, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION: Thomas J.  
; APPLICANT: La Rosa, David K.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 131367  
; LENGTH: 576  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_33439C.1.pcp  
US-10-437-963-131367





Db 126 FLDVVGRIKKISSEIRPANFPDPFKVSDQSLSLRKLDELRLVELKSLEKEGKERVQVME 185  
Qy 186 HNSHLSLCEVLGIDFKQTVYVHPSLDEAGSKNLSNTTIERLAAAANRLKEMKIQRMQ 245  
Db 186 YLKTSLSCVVLGVDPKTISEIHPSLDEAGSPRNISNTTIELMAWAIQRLRETKQRMQ 245  
Qy 246 KLQDFASSMLELWNLMDTPLEEQOMFQNTICNIAAEOEITPNTLSTDFLNVSEVLR 305  
Db 246 KLQDLASTLLELWNLMDTPEEQOAYONITCNIAAEOEITPNTLSTDFLNVSEVLR 298  
Qy 306 LEQKASKMKDLVLKKAELSEHRRRAHLVGBEGYAEBSIEAIEAGIDPSLVLEQIEA 365  
Db 299 ----- 298  
Qy 366 HIATVKEAFSRKDIILEKVERQWQACBEEAWLEDYKNDNRYNAGRGHILTKRAEKART 425  
Db 299 ----- 321  
Qy 426 LVNKIPGMVDVL-----RTKIAAWKNERGKEDFTYDGVSLSSMLDYMFMVROEKEKKRQ 481  
Db 322 LVSKIPGPRQTKPPPRSLQSPRA---IHRXPPALSS-----HATHRSR 367  
Qy 482 RDQKLODQKABOEALYGSKSPSKPLST--KKAPRHSMGANRRLSLGGAT-MOPKPT 538  
Db 368 TNPAPHDPSPTQOTTL-----PVSSPLASHRHLPHPHAFPSPRASYIIRGFTRIQTKT 423  
Qy 539 DILHKSVAARKTEIEIGTLS-----PSSRGLDI-----AGLPIKLSFNA 580  
Db 424 T-----TSSSRSSTTSFSLXVLVSWPCRCFLDMGNVLSCTMAKVGKGARVILPDD 480  
Qy 581 STLRETETPRKPPAQI---TPGNSVSSSTPVRPITNN---TBDDENRTPKTFALNPKTP 633  
Db 481 GGLRQVALPATA-AELMDAPGHFLADARAARVGARLAALSADAELELGAVVATFPMKRL 539  
Qy 634 MVTAPMQAMVPSLAN-----KVSA-----TPVSLVYDKPEVTLQEDID 673  
Db 540 GTPLAPADMARLAATAVREARSAKVAAAVVAPPPTPLQAEDAAPRLRDEMVD 593

## RESULT 10

US-10-425-115-194860

; Sequence 194860, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 194860

; LENGTH: 598

; TYPE: PR

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577\_109303C.1.pep

US-10-425-115-194860

Query Match

Best Local Similarity 34.0%; Score 1187.5; DB 17; Length 598;

Matches 258; Conservative 118; Mismatches 179; Indels 51; Gaps 10;

Qy 11 QMSITCDSLLLELNVWDEVEGPDTRDRMLLEBOECLEVYRRKVDQANRSRQLRKAI 70

Db 14 EAGTSCGALLRELOQIAEVEGSEGEKALLDIERECLVYRRKVDQANRTRVQLHOSV 73

Qy 71 ARGEAELAGICAMGEPVHVVRQSNQKLHGLREELNATVPVLEEMKKKVRNQFVHVI 130

Db 74 AAKEAEVASLMTGLGHEKLYLKK-DKGVVPLKEQLATVAPVLESJKCKEERIKQFSDIR 132

Qy 131 EOIKKIS---SEIRPADVPVKVPVQSDLSLRKLDELTKDLESLOKEKSDRLKQVLEHL 187  
Db 133 SQIEKIRFELSEYNDQDPSLAAEBHDLMSRKLNSYQTLRALQDKSRLRKVLVEYI 192  
Qy 188 NSHLSLCEVLGIDFKQTVYVHPSLDE--AEGSKNLSNTTIERLAAAANRLKEMKIQRMQ 245  
Db 193 NEVHSLCGVLGIDFGSTVHEVHPSLHQNGVQEQRNINSNLTLEGLASTIYKLKAERKSRH 252  
Qy 246 KLQDFASSMLELWNLMDTPLEEQOMFQNTICNIAAEOEITPNTLSTDFLNVSEVLR 305  
Db 253 KAREIMESLCQLMKLMDSPBEKRFQSKVMSSLLPBEGITSPGLSEETIEKMEFEVER 312  
Qy 306 LEQKASKMKDLVLKKAELSEHRRRAHLVGBEGYAEBSIEAIEAGIDPSLVLEQIEA 365  
Db 313 LTRUKTSRLKEIVMKRRAELAIQNAHIEPDVSTAPEQTDALIDSLGIDSELLANIES 372  
Qy 366 HIATVKEAFSRKDIILEKVERQWQACBEEAWLEDYKNDNRYNAGRGHILTKRAEKART 425  
Db 373 QILKAKESLSRKDIMDRINRWIAACDEAWLEBYNQDPKYSAGRGHINLKAERKARI 432  
Qy 426 LVNKIPGMVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDYMFMVROEKEKKRQDOK 485  
Db 433 LVRKIPSMVDDLINRTFAWENARNK-EPLYDGGRLISVLEBYRLSRHQKEENRRYRDOK 491  
Qy 486 KLQDOLKABOEALYGSKSPSKPLS-TKKA---PRHSMGCAN---RRLSLGGATWQPP 536  
Db 492 KLESILLAEKAIPIGSRPSRKTSLSLRKANGVRPNNTNGLKTPTPRRLSLGSATPE-- 549  
Qy 537 KTDILHKSVAARKTEIEIGTLPSSSRGL-----DIAGLPIKLSFNASTL----- 583  
Db 550 -----LLTPRSYSGHNRYFGDVRRLSTSHLNFDDSLSTFTTSIGS 589  
Qy 584 RETETP 589  
Db 590 SEPESP 595

## RESULT 11

US-10-437-963-135276

; Sequence 135276, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 135276

; LENGTH: 589

; TYPE: PR

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_36968C.1.pep

US-10-437-963-135276

Query Match

Best Local Similarity 33.8%; Score 1182.5; DB 16; Length 589;

Matches 255; Conservative 133; Mismatches 176; Indels 31; Gaps 12;

Qy 1 MSSAVKXDLQHLQMTTCDSLLLELNVWDEVEGPDTRDRMLLEBOECLEVYRRKVDQAN 60

Db 1 MSALLRE-----TSCGSLQLKQSVWDEVEGEEDRDVKLFOLDQECLEVDYRKVDQAT 54



Db		74	ADSAEVAALCATIGEPSTTVHACSSLOSTGNLKEELSGITPELEEWRRRRRRRRKFS	133
Qy	3°	128	HVIEIQIKKISSEIRPADFPVKVPVQDSLSRLKDELTKDLESQKSESRKLQVIEHL	187
Db		134	EVTELINRIEQEMKPSK--QHLTMDNSDLTIRLEELRAYLQDLQLEKDSVRVKWTML	191
Qy		188	NSHLSLCEVLGIDFKQTVYEVHPSLDBAEGSKNLSNTTIERLAAAAANLRKMKIORMKL	247
Db		192	GSFHSLSVLGMDFRET--NLH--HDEG--DISDDAIALSVISIGLREIKRNMQKL	244
Qy		248	QDFASSMLELWNLMDTLPBEEQMFQNTICNIAASEQEITERTNTLSTDFLNTYVESEVLRL	307
Db		245	QDLLATWLDLWNLMDTSEBQKRFQSVACNIAASEDEITERDALSMEFINNVAEAVRLE	304
Qy		308	QLKASKMKDLVLKKKAELEBHRRAHLVGBEGYAEBFSEIAEAGAIOPSLVLEQIEAHI	367
Db		305	RLKECRMKDLVLKKYDELNEIRRAHPVENEDDAMMFDAIDSDA--KRSLITERLEVOI	363
Qy		368	ATVKEAFASRKDILEKVERWQNAEEAEWLDEYKDDNRYNAGCAHLTLKBAEKARTLV	427
Db		364	SEAKDEFSRKDVLKKEKQQAALKEEESWLEEEYNNRNNYVNGKTHVLVLKBAEKARLV	423
Qy		428	NKIPGMVDVLTKTAAAMKNSRGKEDFTYDGVLSLSMLDEYMFVROKEQEKKRQDKKL	487
Db		424	SKMPAMAEALITKVIAWEKRGAK--FEYDGGDLGLDMLIEEYNNTRKEQEQRKQDORM	482
Qy		488	QDQLKBAQEALYGSKPSPKPLSTKCAPRHSMGGANRLSLGGATWQPPKTDILHKSRYR	547
Db		483	LGQ-----GTGESP----VVRPPPKNKVTRTSLSGGTSTGGKKA----SASVS	524
Qy		548	AAKKTBEIGTLPSSSRGLD	567
Db		525	SRPSTPSF-LKSPNMSARRSD	543

RESULT 13  
US-10-425-115-287895  
; Sequence 287895, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihui  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 287895  
; LENGTH: 660  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(660)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_25653C.1.pgp  
US-10-425-115-287895

	Query Match	32.6%	Score 1141	DB 17	Length 660
Best Local Similarity	40.3%	Pred. No. 1e-64			
Matches 257	Conservative 115	Mismatches 196	Indels 70	Gaps 12	
15	TCDSLLLEUNVTWDEVGSPD	TTDRMLLELEOECL	EVYRRKVDQANR	SRAQLKAIABGE	74
74	TGSLLLQKLOFTWDEVGSEED	RDKVYLQLOQCL	LDVYKRVQD	ATNSRDLI	133
75	AELAGICSAMGSPPHVHR	QSQKHLGLREELNAI	VPYLEEMKKK	VERWNOFVHVIEQIK	134

Db 134 IELARLLSALGERAI-ARTPEKTTGIIKQOLAAIAPTLTQKQNRKREFVNVQSQID 192  
Qy 135 KISSEIRPADFPVKV---PVDQSDLSLRKLDLTKDLESLOKESDRKQVIEHLSUH 191  
Db 193 QICGEIAGYXVEGEQMTTPQVNBDDTLTLEBEDFRSOLKDLKESXRLEKVLVYGVIVH 252  
Qy 192 SLCEVLGIDFKQVYVHPHSLDEAEGS---KNLSNTTIERIAAANRLREMKIORMOKLOD 249  
Db 253 DLXCVLGMDFSLVIEVHPSLDDSGDNCKSISNDTLTKDNTVATLNEDKKLRLSKQJE 312  
Qy 250 FASSMELNMLMDTPLEEQOMFQNTICNIAAEEQETERTPTLSTDFLNVYVESEVLRLEQ 309  
Db 313 LAGQLVDLMDLMDAPKEERMFHVTNCRSASVDEVTAPGSLALDLIEQAEVEVQRLOD 372  
Qy 310 KASKMKDLVKKAELEEHRRRAHLVGEYAEESIEAIEAGADPSLVLEQIEAHAT 369  
Db 373 KYSMKEIAFKQNELEUEDIYAGAHIVIDTAAAHKILALIEAGNTEPSELIADMDAQIAK 432  
Qy 370 VKEEAFSRKDIILEKVERWQNAEEBAWLEDYKNDNRVYNAAGRAHLTLKRAEKARTLVNK 429  
Db 433 AKEEALSRLDKLVKVERWWSACEEESWLEDYKNDNRVYNSRGAHLNLERAEAGEGEXSK 492  
Qy 430 -----IPGWDVLRTKIAANKNERGKEDFTYDGVLSMLDEYMFVROEKEKKRQD 483  
Db 493 DSAKTAPGNSAATAAPVGL--TATAKED-----AALLAMLDYVYMLRQEREKEKRWRE 545  
Qy 484 OKK-LQDQKARQEAALYKSPSPKPLSTKKAAPRHSWGGA-----NRRLSLGATWQP 535  
Db 546 QKRYVEQQLNTHDEGPFGRVFNRPASSKKAIGPKLNGSVNGTTPNNRUSISGQ---- 601  
Qy 536 PKTDILHKSVRAAKTEBIGTLPSSSRGLDIAPLPIKLSFNASTLRLETETPRKPPAQ 595  
Db 602 -----QNGGGHGVRSKGKDKK-----DTAK-----T 623  
Qy 596 ITPGNSVSTPVRPINTNEDDENRTPKFTALNPKTP 633  
Db 624 ASFGNSAAATAPVGLTATAKEDA-ASQISVTDLVPESTP 660

## RESULT 14

US-10-437-963-106958  
; Sequence 106958, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 106958  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_11356C.1.pap  
US-10-437-963-106958

Query Match 32.3%; Score 1127.5; DB 16; Length 610;  
Best Local Similarity 41.4%; Pred. No. 6.6e-64;  
Matches 260; Conservative 107; Mismatches 212; Indels 49; Gaps 14;  
Qy 1 MSSAVKQDLQHQWSTTCDSLLLELVNVDVEGPDTRDRMLLEOECELVYVRKVDQAN 60  
Db 11 MDKAVRSV-SFETPCGRLLELEQIWIETGEIEREDKDMFLETECMVRVRRKVDSAN 69

Qy 61 RSRAQLRKAIAEGEAEAGLAGICSAMGE--PPVHVQSNQKHLGLREELNAIVPYLEEMKKK 118  
Db 70 AERSQLQSSLMAXEAEALKVLVASIGETTPKPKVDEK-----QSLKEQLAKVTPLEDLRSK 125  
Qy 119 KVERWNPQVHVIEQIKKISSEI---RPADVPVPKVPVQSDLSLRKLDLTKDLESLOK 174  
Db 126 KEERIKQFSLVQSQIEKIKAQISDHNQHDGPNVHSDKNDHDLSTRRLSDLOAERNLQK 185  
Qy 175 EKSRLQVIEHLSHLSLCEVLGIDFKQTVYVHPHSLDEA--EGSKNLSNTTIERIAA 232  
Db 186 EKSRLQVITYVDEVHCLCSVLGMDFAKTVKDVHPSLHGANSSENSTNISDSTLEGLTET 245  
Qy 233 ANRLREMKIORMOKLODFASMLELWNLMDTPLEEQOMFQNTICNIAAEEQETERTPTLS 292  
Db 246 ILKLAERKTRVSKLOBIQVGLKHLNLMESTEQERRHFTRVAAVLGSTEEITSSVLS 305  
Qy 293 TDFLNVYSEVLRLEQLKASKMKDLVLKKAELAEHRRRAHLVGEYAEESIEAIEAG 352  
Db 306 LETTQETEEVERLTQKASRMKELVLKKELEEDICSNAMPEMDSTAEKITALIDSG 365  
Qy 353 AIDPSLVLEQIEAHATVKEEAFSRKDIILEKVERWQNAEEBAWLEDYKNDNRVYNAAGR 412  
Db 366 LVDPCELLSSITETQIAKARBSLTKDIMEKVDRLWSACDEETWLEBYNQDSSRSYAGR 425  
Qy 413 AHLTLKRAEKARTLVNKIPGWDVLRTKIAANKNERGKEDFTYDGVLSMLDEYMFVRQ 472  
Db 426 AHINLKRAEKARILVQKIPSMIDNLIAKTFAWEDER-SVPFLYDGLARLVAILEEQKLRV 484  
Qy 473 EKEKKRQRDQKQDLQKAEQEAALYKSPSPKPLS-TKKAPRHSWGGAANRRLSLGGA 531  
Db 485 QKEEDKRRHRDQKQLQSLLLKKELEIFGSKPKTSSFNRTSSHPNG-----NGA 537  
Qy 532 TMQPPKTDILHKSVRAAKTEBIGTLPSSSRG-----LDIAGLPIKLSFNAST 582  
Db 538 GFMTVP-----PRVSAGSATPEL--LTPRSYSGRYNNYFKENRRLAAAPL-----NFST 585  
Qy 583 LRETETPRKPPAQITPGNSVSTPVRPI 610  
Db 586 VSK-EDSMSSSFASIS-----GSEPDSP 607

## RESULT 15

US-10-425-115-264132  
; Sequence 264132, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 264132  
; LENGTH: 611  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_172503C.1.pap  
US-10-425-115-264132

Query Match 31.5%; Score 1101.5; DB 17; Length 611;  
Best Local Similarity 41.1%; Pred. No. 3.1e-62;  
Matches 245; Conservative 113; Mismatches 207; Indels 31; Gaps 9;  
Qy 1 MSSAVKQDLQHQWSTTCDSLLLELVNVDVEGPDTRDRMLLEOECELVYVRKVDQAN 60  
Db 11 MDKALRASV-SFTTPCGALLRELEQIWIETGEIEREQDKDMFLETECMVRVRRKVDSAN 69

Search completed: November 20, 2004, 00:44:43  
Job time : 89 secs

**THIS PAGE LEFT BLANK**